Hints against the cold and collisionless nature of dark matter from the galaxy velocity function

Schneider, Aurel ; Trujillo-Gomez, Sebastian ; Papastergis, Emmanouil ; Reed, Darren S ; Lake, George

Abstract: The observed number of dwarf galaxies as a function of rotation velocity is significantly smaller than predicted by the standard model of cosmology. This discrepancy cannot be simply solved by assuming strong baryonic feedback processes, since they would violate the observed relation between maximum circular velocity (vmax) and baryon mass of galaxies. A speculative but tantalizing possibility is that the mismatch between observation and theory points towards the existence of non-cold or non-collisionless dark matter (DM). In this paper, we investigate the effects of warm (WDM), mixed (MDM, i.e. warm plus cold), and self-interacting DM (SIDM) scenarios on the abundance of dwarf galaxies and the relation between observed H i line width and maximum circular velocity. Both effects have the potential to alleviate the apparent mismatch between the observed and theoretical abundance of galaxies as a function of vmax. For the case of WDM and MDM, we show that the discrepancy disappears, even for lukewarm models that evade stringent bounds from the Lyman- forest. SIDM scenarios can also provide a solution as long as they lead to extended (1.5 kpc) DM cores in the density profiles of dwarf galaxies. Only models with velocity-dependent cross-sections can yield such cores without violating other observational constraints at larger scales.

DOI: https://doi.org/10.1093/mnras/stx1294

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **470,** 1542–1558 (2017)
Advance Access publication 2017 May 29

doi:10.1093/mnras/stx1294

Hints against the cold and collisionless nature of dark matter from the galaxy velocity function

Aurel Schneider,[1]★ Sebastian Trujillo-Gomez,[2] Emmanouil Papastergis,[3] Darren S. Reed[2,4] and George Lake[2]

[1]*Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*
[2]*Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*
[3]*Kapteyn Astronomical Institute, University of Groningen, Landleven 12, Groningen NL-9747AD, The Netherlands*
[4]S[3]IT, *University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*

Accepted 2017 May 23. Received 2017 May 22; in original form 2016 December 21

ABSTRACT

The observed number of dwarf galaxies as a function of rotation velocity is significantly smaller than predicted by the standard model of cosmology. This discrepancy cannot be simply solved by assuming strong baryonic feedback processes, since they would violate the observed relation between maximum circular velocity ($v_{\rm max}$) and baryon mass of galaxies. A speculative but tantalizing possibility is that the mismatch between observation and theory points towards the existence of non-cold or non-collisionless dark matter (DM). In this paper, we investigate the effects of warm (WDM), mixed (MDM, i.e. warm plus cold), and self-interacting DM (SIDM) scenarios on the abundance of dwarf galaxies and the relation between observed H I line width and maximum circular velocity. Both effects have the potential to alleviate the apparent mismatch between the observed and theoretical abundance of galaxies as a function of $v_{\rm max}$. For the case of WDM and MDM, we show that the discrepancy disappears, even for lukewarm models that evade stringent bounds from the Lyman-α forest. SIDM scenarios can also provide a solution as long as they lead to extended ($\gtrsim$1.5 kpc) DM cores in the density profiles of dwarf galaxies. Only models with velocity-dependent cross-sections can yield such cores without violating other observational constraints at larger scales.

Key words: galaxies: dwarf – cosmology: theory – dark matter.

1 INTRODUCTION

Observations of the abundance and structure of dwarf galaxies have the potential to probe the particle nature of dark matter (DM). This is because effects from DM free-streaming or from (self-) interactions have an impact on structure formation at the smallest observable scales.

There are potential inconsistencies between small-scale observations and the standard model of Λcold dark matter (ΛCDM, based on the observations of the *Planck* satellite, Planck Collaboration XXIV 2016). First of all, haloes predicted by gravity-only simulations greatly outnumber observed galaxies. This long-standing discrepancy has been established for both the Milky Way (MW) satellites (Klypin et al. 1999; Moore et al. 1999) as well as nearby isolated galaxies (e.g. Tikhonov et al. 2009; Zavala et al. 2009) and is usually referred to as the *overabundance* (or *missing satellite*) problem. Second, observations of rotation velocities from stars and gas point towards very shallow inner density profiles of small haloes in strong contrast to predictions from gravity-only simulations. This is generally known as the *cusp-core* (e.g. de Blok 2010) or the *too-big-to-fail* problem (TBTF, Boylan-Kolchin, Bullock & Kaplinghat 2011; Papastergis et al. 2015), depending on the context.

The main difficulty with these problems of small-scale structure formation is the fact that they are based on predictions from gravity-only simulations, ignoring any potential effects from baryonic physics. Indeed, it is expected that photoevaporation from ultraviolet (UV) sources during reionization can expel gas from small haloes, effectively preventing star formation and reducing the number of observable dwarf galaxies (e.g. Gnedin 2000; Okamoto, Gao & Theuns 2008). More recently, it was realized that supernova feedback is energetic enough to reshape the inner parts of halo profiles making them considerably shallower (Governato et al. 2012). However, the details of how feedback affects the halo profile are still under debate. For example, Onorbe et al. (2015) and Read, Agertz & Collins (2016b) show that the core size depends on the details of the star formation history, while Di Cintio et al. (2014) and Fitts et al. (2016) connect the core size to the dwarf's stellar-to-halo mass ratio. Other papers (e.g. Sawala et al. 2016; Fattahi et al. 2016) point out that the presence of cores depends on the feedback implementation and might not be required to recover the observations.

★E-mail: aurel.schneider@phys.ethz.ch

Despite the ongoing debate about the efficiency of baryon-induced feedback mechanisms, they are generally assumed to be the most likely explanation for the dwarf abundance and structure problems. However, because of our poor understanding of subgrid effects in hydrodynamical simulations, it has so far been impossible to verify these assumptions from first principles.

A very useful statistic that simultaneously probes both the abundance and the structure of galaxies is the velocity function (VF), i.e. the number density of galaxies as a function of their observed rotation velocity. The VF offers a direct link between observations and theory because the rotation velocities of galaxies act as a tracer of the halo gravitational potential. Since the galaxy VF is sensitive to both the abundance and the inner structure of haloes, any model that predicts the observed VF is also likely to solve the overabundance, the cusp-core, and the TBTF problems. This makes the VF an ideal probe of structure formation.

In a recent study, Klypin et al. (2015) compiled the VF of the local 10 Mpc around the MW using rotation velocities ($v_{\rm rot}$) predominantly based on spatially unresolved H I line widths. When compared to the VF from gravity-only simulations, they found a discrepancy in the abundance of galaxies at $v_{\rm rot} < 80\,{\rm km\,s^{-1}}$ which increases towards lower velocities. However, this comparison is based on the assumption that $v_{\rm rot}$ is a good approximation of the maximum circular velocity ($v_{\rm max}$) of the halo, which is not guaranteed.

In a recent paper, we investigated the relation between the measured rotation velocity of a galaxy and the maximum circular velocity of the halo that hosts the galaxy, and we studied how the resulting $v_{\rm max}$ based VF is affected by baryonic processes (Trujillo-Gomez et al. 2016, hereafter TG16). Our main findings were the following: (i) the bias between $v_{\rm rot}$ and $v_{\rm max}$ is not large enough to significantly reduce the mismatch between the observed VF and the one predicted by gravity-only simulations of ΛCDM; (ii) while baryonic processes are able to reduce the theoretical abundance of galaxies and alleviate the overabundance problem, they cannot completely solve it without simultaneously violating the observed relation between baryon mass and $v_{\rm max}$ [i.e. the $v_{\rm max}$ baryonic Tully–Fisher (BTF) relation].

This paper builds upon TG16 and investigates the tantalizing possibility that the mismatch between the observed and predicted galaxy VF is caused by the underlying particle properties of DM. Both non-cold and non-collisionless DM models could provide more natural solutions, as they suppress the amplitude of matter perturbations and/or alter the halo density profiles. As representative examples, we focus on the effective scenarios of warm (WDM), mixed (MDM, i.e. warm plus cold), and self-interacting DM (SIDM).

The paper is structured as follows: in Sections 2 and 3, we give a brief summary of the results obtained in TG16, and we present our theoretical model of the DM halo VF based on the extended Press–Schechter (EPS) approach. Sections 4–6 present the results for WDM, MDM, and SIDM. We examine how these DM scenarios affect the halo profiles, the maximum circular velocities, and finally the VF. In Section 7, we discuss qualitatively other potential DM particle scenarios. Our results are summarized in Section 8.

2 SETTING UP THE PROBLEM

The number density of galaxies as a function of rotational velocities – i.e. the VF – is a very useful observational quantity relating information about galaxy abundance with the underlying halo potentials. This allows to compare theory with observations without detailed knowledge about the galaxy formation efficiency. In this section, we recap the results from TG16, summarizing the procedure for obtaining maximum circular velocities ($v_{\rm max}$) from observed H I line widths (w_{50}) and how this affects the shape of the VF.

2.1 Galaxies in the local universe

In a recent paper, Klypin et al. (2015) performed a detailed analysis of the abundance of galaxies within the local volume around the MW. Their analysis is based on the galaxy catalogue from Karachentsev, Makarov & Kaisina (2013, hereafter K13) which they show to be complete down to a limiting magnitude of $M_{\rm B} = -12$ within 10 Mpc from the MW. The rotation velocities ($v_{\rm rot}$) of all galaxies in the K13 sample were determined by either relying on inclination corrected unresolved H I line-width measurements or by using the magnitude–velocity relation for the galaxies with no detected H I (a fraction of less than 10 percent). Klypin et al. (2015) find a slowly rising VF down to $v_{\rm rot} \sim 15\,{\rm km\,s^{-1}}$ which they claim to be in tension with the ΛCDM model prediction below $v_{\rm rot} \sim 80\,{\rm km\,s^{-1}}$. Similar conclusions have been made previously by Zavala et al. (2009), Trujillo-Gomez et al. (2011), and Papastergis et al. (2011) based on data from the HIPASS and ALFALFA H I surveys, respectively.[1]

The statement that there is tension between observations and the ΛCDM model relies on the assumption that the rotational velocity from H I line widths, $v_{\rm rot} \equiv w_{50}/(2\sin i)$, can be used as a proxy for the maximum circular velocity ($v_{\rm max}$) of a halo. Klypin et al. (2015) showed that this is approximately the case, at least for a subset of selected dwarf galaxies and assuming all haloes to have Navarro–Frenk–White (NFW) profiles with a given concentration–mass relation. More recently, several authors have questioned the validity of these assumptions, reporting strong biases between $v_{\rm rot}$ and $v_{\rm max}$ instead. These studies used abundance matching (AM, Brook & Di Cintio 2015b; Brook & Shankar 2016), measurements from zoom-in hydrodynamical simulations of individual galaxies (Macciò et al. 2016), or semi-analytical models (Yaryura et al. 2016; Obreschkow et al. 2013).

In TG16, we used direct observations of galaxies to perform a detailed investigation of possible biases between $v_{\rm rot}$ and $v_{\rm max}$. We showed that $v_{\rm max}$ can be directly recovered from $v_{\rm rot}$, and we argued that the bias between the two is present but not large enough to solve the discrepancy present in the VF. We will now summarize the method developed in TG16.

2.2 From $v_{\rm rot}$ to $v_{\rm max}$

Most of the galaxies from the K13 sample only have H I line-width measurements (or magnitude-based estimates for the subdominant population of gas-free galaxies) without any spatial information, making it impossible to estimate the corresponding $v_{\rm max}$. There is, however, more information for a subsample of galaxies with existing spatially resolved measurements of their kinematics. A sample of 200 galaxies with interferometric H I observations ($v_{\rm out}$) at the outermost H I radius ($r_{\rm out}$) was compiled by Papastergis & Shankar (2016). We use this catalogue to select all galaxies with $r_{\rm out} > 3r_{1/2}$,

[1] Recently, Bekeraitė et al. (2016) have shown that there is disagreement between observations and simulations at larger velocities as well (i.e. between 60 and 300 km s^{-1}). This tension is, however, not as strong and could be due to the fact that $v_{\rm max}$ of larger galaxies is dominated by the stellar component.

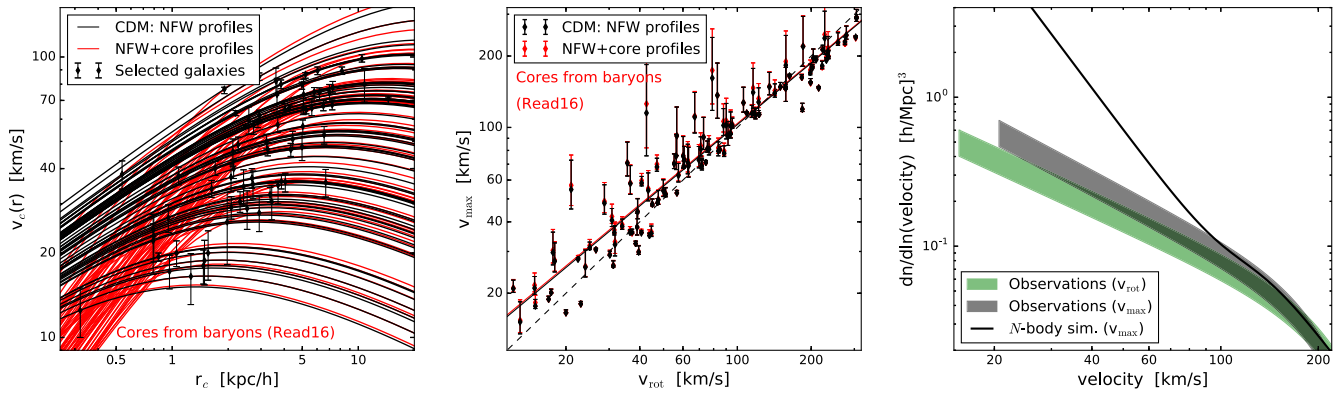

Figure 1. *Left:* observed circular velocities ($v_{\rm out}$ at $r_{\rm out}$) from the selected galaxy sample (black symbols) together with fitted velocity profiles based on NFW (black lines) and NFW+core (red lines, accounting for baryon-induced cores). *Middle:* relation between $v_{\rm rot}$ of galaxies and $v_{\rm max}$ of haloes from NFW fits (black symbols) and NFW+core fits (red symbols) with error bars indicating the change of $v_{\rm max}$ when the halo concentrations are varied by 1σ around the mean. Corresponding linear fits are shown as black and red lines. The effects from baryon-induced cores are very small because $r_{\rm out}$ lies well beyond the core radii. *Right:* VF based on $v_{\rm rot}$ (green band, from Klypin et al. 2015) and $v_{\rm max}$ (grey band, including $v_{\rm max}$ correction from the middle panel) together with the prediction from gravity-only simulations. The rotation velocity ($v_{\rm rot}$) is defined via H I line width (w_{50}) and galaxy inclination (i), i.e. $v_{\rm rot} \equiv w_{50}/(2\sin i)$.

where $r_{1/2}$ is the galactic half-light radius (see TG16). This additional selection criterion guarantees that the velocity measurement is not dominated by baryonic effects, including a potential DM core from strong stellar feedback (Read et al. 2016b). The final catalogue consists of 109 galaxies (distributed over the full range of relevant scales) which can be used to estimate the relation between $v_{\rm rot}$ and $v_{\rm max}$.

In a first step, we fit NFW profiles (Navarro, Frenk & White 1996) to the observed velocities $v_{\rm out}$ at radius $r_{\rm out}$ in order to determine the corresponding $v_{\rm max}$, which we then compare to $v_{\rm rot}$ from H I line-width measurements.[2] For the concentrations, we use the relation from Dutton & Macciò (2014) based on *Planck* cosmology.

In the left-hand panel of Fig. 1, the spatially resolved observed circular velocities $v_{\rm out}$ (black symbols with error bars) are shown together with the fitted velocity profiles based on NFW (grey lines). For most galaxies in the sample, $v_{\rm max}$ is not much larger than $v_{\rm out}$, owing to the fact that $v_{\rm out}$ is observed far out in radius ($r_{\rm out}$). The relation between $v_{\rm rot}$ (i.e. the rotation velocity, obtained via the measurement of unresolved H I line widths) and $v_{\rm max}$ is shown in the middle panel of Fig. 1 (black symbols). The error bars illustrate the sensitivity of the results for varying concentrations within the 1σ scatter given by Dutton & Macciò (2014). In general, there is a small bias between $v_{\rm rot}$ and $v_{\rm max}$ slowly growing towards very small velocities.

In a second step, we attempt to include effects from baryons on the circular velocity profiles. We therefore repeat the same analysis using the fit from Read et al. (2016b, herafter R16) which consists of an NFW profile plus a baryon-induced core proportional to the stellar half-light radius. The mass profile of the R16 fit is given by

$$M_{\rm R16}(r) = M_{\rm nfw}(r)\, f^n, \quad f = \left[\tanh\left(\frac{r}{r_{\rm c}}\right)\right] \qquad (1)$$

and is a simple extension of the NFW mass profile ($M_{\rm nfw}$) with two additional free parameters $r_{\rm c}$ and n. R16 showed that the sizes of baryon-induced cores are proportional to the half-light radii (i.e. $r_{\rm c} = \eta\, r_{1/2}$) and that n varies between 0 and 1 depending on the individual star formation history of each galaxy. Here, we fix $\eta = 1.75$ because this was shown by R16 to provide the best match to their simulations. Furthermore, we adopted the value $n = 1$ for the second free parameter in order to maximize the effect from baryons.

The effects of the baryon-induced cores are illustrated in the left-hand and middle panels of Fig. 1. The inner part of the velocity profiles are much steeper (left-hand panel, red lines) reflecting how cores affect the velocity profiles. However, the resulting values for $v_{\rm max}$ (middle panel, red symbols) are nearly indistinguishable from the ones obtained with an NFW fit. This shows that baryon-induced cores do not bias the $v_{\rm max}$ estimates of the selected galaxy sample, which is a direct result of our original selection criteria (i.e. $r_{\rm out} > 3r_{1/2}$). Such a selection greatly simplifies the analysis and can be justified as long as the resulting galaxy sample is *representative* for the galaxies that make up the VF. In TG16, we used the BTF relation to show that this is indeed the case.

2.3 Corrected velocity function

Let us now turn our attention to the VF and how it can be corrected to account for the bias between the rotational ($v_{\rm rot}$) and the maximum circular velocity ($v_{\rm max}$). This correction is important, since only $v_{\rm max}$ can be directly related to the halo mass (and therefore to the theory prediction), while $v_{\rm rot}$ depends on the details of the gas distribution within the galaxy.

In the right-hand panel of Fig. 1, we plot both the observed VF based on $v_{\rm rot}$ (green band) and the predicted VF from gravity-only simulations[3] of a ΛCDM universe based on $v_{\rm max}$ (black solid line). The two VFs agree reasonably well at large velocities but start to diverge below 100 km s^{-1}. This apparent discrepancy between observations and theory has been pointed out repeatedly in the past (see e.g. Tikhonov et al. 2009; Zavala et al. 2009; Trujillo-Gomez et al. 2011; Papastergis et al. 2011).

It is possible to correct the observed VF using the average relation between $v_{\rm max}$ and $v_{\rm rot}$ obtained with the selected galaxy sample

[2] All observed values of $v_{\rm out}$ are corrected for pressure support (see section 4.1 in Papastergis & Ponomareva 2016). For large galaxies (with $v_{\rm out} > 120$ km s^{-1}), we furthermore subtract the expected contribution from stars and cold gas in the galaxy centres. We have checked that this correction does not affect our final results (see TG16 for more details).

[3] The line is based on the MultiDark suite of simulations (Klypin et al. 2016) and includes a correction for the increase of circular velocities due to the stellar component of galaxies visible beyond 80 km s^{-1} (see Trujillo-Gomez et al. 2011).

(i.e. the fit from the middle panel of Fig. 1). The resulting $v_{\rm max}$-corrected VF is plotted as a grey band in the right-hand panel of Fig. 1. Despite being slightly steeper, it remains inconsistent with the prediction from gravity-only simulations. In the following, we use this $v_{\rm max}$-corrected VF and compare it to theoretical predictions including baryon effects as well as modifications induced by the DM model.

3 THEORETICAL PREDICTIONS FOR THE VELOCITY FUNCTION OF HALOES

The VF is sensitive to both baryonic feedback effects and the particle nature of DM, making an accurate modelling both essential and challenging. In this paper, we use an analytical approach based on the EPS model. This has the advantage of being easily adaptable to different DM models, and it can be used to estimate suppression effects from baryons.

3.1 Modelling the velocity function

The calculation of the halo VF is based on the EPS approach presented in Schneider, Smith & Reed (2013) and Schneider et al. (2014). The first and most important step is to obtain the halo mass function with sharp-k filter

$$\frac{{\rm d}n}{{\rm d}\ln M} = \frac{1}{12\pi^2}\frac{\bar{\rho}}{M}\nu f(\nu)\frac{P_{\rm lin}(1/R)}{\delta_{\rm c}^2 R^3}, \tag{2}$$

$$\sigma^2(R) = \int \frac{{\rm d}\mathbf{k}^3}{(2\pi)^3} P_{\rm lin}(k)\Theta(1-kR), \tag{3}$$

where $P_{\rm lin}(k)$ is the linear power spectrum, $\delta_{\rm c} = 1.686$ the collapse threshold, and Θ the Heaviside step function. The first crossing distribution $f(\nu)$ is obtained from the ellipsoidal collapse model, yielding

$$f(\nu) = A\sqrt{2\nu/\pi}(1+\nu^{-p}){\rm e}^{-\nu/2} \tag{4}$$

with $\nu = (\delta_{\rm c}/\sigma)^2$, $A = 0.322$, and $p = 0.3$. The halo mass is assigned to the filter scale with the relation $M = 4\pi\bar{\rho}(cR)^3/3$ where $c = 2.5$ (see also Benson et al. 2013, for a similar description).

In order to obtain the maximum circular velocity, we assume all haloes to be described by an NFW profile. This is a good assumption, even for alternative DM (ADM) models with extended cores, because the radius corresponding to $v_{\rm max}$ lies beyond the scale radius of the halo. For the concentration–mass relation of the ΛCDM model, we use the power-law relation

$$c(M) = 10^{1.025}\left(\frac{10^{12}\,{\rm M}_\odot\,h^{-1}}{M}\right)^{0.097} \tag{5}$$

from Dutton & Macciò (2014) based on the *Planck* cosmology. ADM scenarios can have different concentrations and we follow the approach from Schneider (2015) which consists of comparing halo formation times between CDM and ADM models and assigning concentrations accordingly. An estimate of the average redshift of halo formation can be obtained by solving the equation

$$D(z_{\rm c}) = \left[1+\sqrt{\frac{\pi}{2}}\frac{1}{\delta_{\rm c}}\sqrt{\sigma^2(F^{1/3}R)-\sigma^2(R)}\right]^{-1} \equiv \frac{5\Omega_{\rm m}}{2}H(z_{\rm c})\int_{z_{\rm c}}^{\infty} dz\frac{(1+z)}{H(z)^3} \tag{6}$$

for the collapse redshift $z_{\rm c}(M)$, where $F = 0.05$. Once the function $z_{\rm c}(M)$ is known for both ADM and CDM, we can link together ADM and CDM haloes with the same collapse redshift and assign concentrations for ADM haloes from equation (5). Although $z_{\rm c}(M)$ is a rather poor estimate of the *actual* collapse redshift measured in simulations, the resulting concentrations of ADM haloes are surprisingly accurate (see Schneider 2015).

To finally obtain the VF, we create a mock sample of haloes drawn from the halo mass function, and we assign concentrations from a lognormal distribution. This allows us to determine a value for the maximum circular velocity according to the relation

$$v_{\rm max} = 0.465\sqrt{\frac{GM}{r_{\rm vir}}}\left[c^{-1}\ln(1+c)-(1+c)^{-1}\right]^{-1/2} \tag{7}$$

(directly resulting from the NFW profile, see Sigad et al. 2000) and to re-bin the sample in order to obtain the VF

$$\Phi(v_{\rm max}) \equiv \frac{{\rm d}n}{{\rm d}\ln v_{\rm max}}. \tag{8}$$

Similar approaches have been applied by several authors in the past (see e.g. Zavala et al. 2009; Schneider et al. 2014).

It was shown by Trujillo-Gomez et al. (2011) and Dutton et al. (2011) that the maximum circular velocity of small galaxies are not affected by baryonic infall or contraction. Larger galaxies with $v_{\rm max} \gtrsim 100\,{\rm km\,s^{-1}}$ have boosted velocities due to their baryonic components. We follow Klypin et al. (2015) and correct the maximum circular velocity of massive galaxies by solving the equation

$$v_{\rm max}^{\rm dmo} = v_{\rm max}\left[1+\frac{0.35(v_{\rm max}/120\,{\rm kms^{-1}})^6}{1+(v_{\rm max}/120\,{\rm kms^{-1}})^6}\right]^{-1}, \tag{9}$$

where $v_{\rm max}^{\rm dmo}$ stands for the maximum circular velocity without baryonic correction.

The EPS approach with sharp-k filter has two distinctive advantages with respect to other methods: first of all, it accurately describes the halo abundance of models with arbitrary power spectra, while the standard EPS model with a real-space tophat filter only works for the CDM scenario (Schneider 2015). Second, it does not suffer from artificial clumping, which is a serious problem for direct simulations of DM scenarios with suppressed power spectra (see e.g. Wang & White 2007; Lovell et al. 2014; Hahn & Angulo 2016; Hobbs et al. 2016).

One drawback of the EPS approach is that it does not account for substructures. To correct for this, we multiply the EPS VF by a factor of 1.25 so that it matches the predictions from the MultiDark N-body simulations (Klypin et al. 2016). This corresponds to adding a constant number of subhaloes to each velocity bin.[4] The normalization is done once and is not changed for different DM models.

3.2 Maximizing effects from baryons

The great majority of work on the VF has been based on gravity-only N-body simulations in the past (e.g. Gonzalez et al. 2000; Zavala et al. 2009; Zwaan, Meyer & Staveley-Smith 2010; Papastergis et al. 2011; Obreschkow et al. 2013; Klypin et al. 2015). There are, however, two distinct effects from baryons which should be accounted for, since they have the potential to significantly alter the VF at dwarf galaxy scales. The first effect is *baryonic depletion*

[4] A constant ratio of subhalo to host-halo numbers is a very good approximation for haloes with $v_{\rm max} < 150\,{\rm km\,s^{-1}}$. In Klypin, Trujillo-Gomez & Primack (2011), it was shown that this ratio does not change by more than four percent in the range $v_{\rm max} = 30$–$150\,{\rm km\,s^{-1}}$.

and consists of a reduction of the maximum circular velocity due to the fact that some of the gas is being pushed out of haloes, reducing the total mass and accretion rate of the halo during its formation. The second effect is *baryonic suppression*, referring to the fact that feedback can reduce the number of observable galaxies by pushing the luminosity below the sensitivity level of a given survey. The maximum suppression of the VF from both types of baryonic effects was quantified in TG16. Here, we summarize these results and show how they can be extended to ADM scenarios.

The maximum effect from *baryonic depletion* can be obtained by calculating the VF for a cosmology where the entire baryon content is removed. This is achieved by replacing $\sigma_8 \rightarrow (1 - \Omega_b/\Omega_m)\sigma_8$ as well as $\Omega_m \rightarrow (1 - \Omega_b/\Omega_m)\Omega_m$, resulting in a scale-independent decrease of the maximum circular velocities, i.e.

$$v_{\rm max}^{\rm depl} \simeq 0.86\, v_{\rm max}, \tag{10}$$

independent of the DM model (see TG16 for more details). We want to stress that this corresponds to the *maximum* baryonic depletion, likely to overestimate the true effect.

The second type of effect, the *baryonic suppression*, is more difficult to model as it crucially depends on the details of the suppression mechanism. In TG16, we developed a model-independent approach to quantify the maximum possible suppression of dwarf galaxy numbers. Any decrease of the stellar or gaseous content of galaxies leads to a bend in the relation between $v_{\rm max}$ and $M_{\rm bar}$ – the BTF relation of $v_{\rm max}$. The maximum allowed suppression can therefore be directly constrained by the data without prior knowledge of feedback mechanisms.[5]

In this paper, we describe the $v_{\rm max}$–$M_{\rm bar}$ relation with the function $\mathcal{M}(v_{\rm max})$ which provides a good description of the data points. The suppression induced by potential baryonic processes is furthermore parametrized as

$$\mathcal{M}_{\rm supp}(v_{\rm max}) = [1 + (v_s/v_{\rm max})^4]^{-5}\mathcal{M}(v_{\rm max}) \tag{11}$$

where v_s is a free model parameter. This function leads to a very similar suppression than the one obtained in hydro simulations (see e.g. Sales et al. 2017). A more general parametrization is discussed in TG16.

For the fiducial case of CDM, the function $\mathcal{M}(v_{\rm max})$ is given by a linear least-squares fit of the data. It can be shown that this is indeed a very good fit to the data (see TG16). Next, we can use the *likelihood ratio analysis*[6] to determine the model $\mathcal{M}_{\rm supp}(v_{\rm max})$ with the largest value of v_c that is still in agreement with the data at 3σ confidence level (CL). This function is defined as the model with maximum allowed baryon suppression. For CDM, it is given by the parameter $v_s = 23\,{\rm km\,s^{-1}}$.

In the right-hand panel of Fig. 3, the $v_{\rm max}$–$M_{\rm bar}$ relation of CDM (empty triangles) is plotted together with the linear fit $\mathcal{M}$ (solid black line) and the function of maximum allowed baryon suppression $\mathcal{M}_{\rm supp}$ (dashed black line). While the former is a good fit to the data, the latter is characterized by a strong downturn towards small velocities.

[5] The argument that baryonic feedback effects suppress the BTF relation is only true if $v_{\rm max}$ is used for the velocity, as it is a direct measure of the halo potential and does not depend on the extent of observable gas (as is the case for $v_{\rm rot}$ for example).

[6] The logarithmic likelihood ratio test is based on the measure $D \equiv 2\ln(\mathcal{L}_m/\mathcal{L}_0)$, where $\mathcal{L}_m$ is the maximum likelihood and $\mathcal{L}_0$ the likelihood of a constrained model with fixed v_s. For a large sample size, D is known to be χ^2-distributed and a model with given v_s can therefore be excluded at the CL given by the p-value from a χ^2 statistic.

For the ADM models discussed in this paper, the values of $v_{\rm max}$ are modified with respect to CDM. As a consequence, the $v_{\rm max}$–$M_{\rm bar}$ relation cannot be described by a linear least-squares fit anymore. A more accurate model for $\mathcal{M}$ is obtained when the linear fit from CDM is corrected by accounting for the difference in the average value of $v_{\rm max}$ between ADM and CDM. Based on this corrected function $\mathcal{M}$ for ADM, the maximum baryon suppression model can again be obtained with a likelihood ratio analysis.

Any downturn in the $v_{\rm max}$–$M_{\rm bar}$ relation is expected to have an influence on the VF, as it sets the velocity scale below which galaxies become undetectable by a given survey. At the velocity scale v_c, where $\mathcal{M}_{\rm supp}(v_{\rm max})$ crosses the survey detectability limit (in terms of baryon mass), half of the galaxies are too faint to be visible in the VF. This effect can be modelled as follows

$$\Phi_{\rm supp}(v_{\rm max}) = \mathcal{G}_{\rm supp}(v_{\rm max})\Phi(v_{\rm max}),$$
$$\mathcal{G}_{\rm supp}(v_{\rm max}) = \frac{1}{2}\left[{\rm erf}\left(\frac{\log v_{\rm max} - \log v_c}{\sqrt{2}\log\sigma_c}\right) + 1\right], \tag{12}$$

where we assumed a lognormal distribution of galaxies around the mean (see TG16 for more details). For the K13 sample, the detectability limit is at $\sim 4.3 \times 10^6\ {\rm M}_\odot\, h^{-1}$ resulting in $v_c \sim 29.5\,{\rm km\,s^{-1}}$ for CDM (as indicated by the black cross in the right-hand panel of Fig. 3). The scatter can be directly measured from the data (ignoring baryon suppression) resulting in $\log\sigma_c \sim 0.15$ for most models (including CDM), except for SIDM where the scatter is larger (see Section 6).

In the following sections, the theoretical model developed here is applied to CDM, WDM, MDM, and SIDM models. We always present both the VF without baryon effects as well as the VF with maximum baryon suppression and depletion. These two extreme models quantify the current uncertainty of theory predictions due to unknown feedback effects.

4 WARM DARK MATTER

The first alternative paradigm we consider is the collisionless WDM model, which is characterized by a steep suppression of the power spectrum at small scales caused by the free-streaming properties of the DM particles. The scale and exact shape of the suppression depends on both the DM particle mass and its phase-space distribution. In this section, we restrict ourselves to the standard and most studied case of a Fermi–Dirac distributed DM fluid (i.e. the so-called thermal WDM) for which we vary the particle mass. However, other distributions are possible depending on the exact DM production mechanism (see e.g. Merle & Schneider 2015). These models might lead to somewhat shallower suppressions of the power spectrum, closer to the case of MDM (see Section 5).

There are various constraints of the thermal WDM scenario from the Lyman-α forest (Seljak et al. 2006; Viel et al. 2005, 2013; Baur et al. 2016), the dwarf galaxy abundance in the local volume (Polisensky & Ricotti 2011; Kennedy et al. 2014; Horiuchi et al. 2014), high-redshift galaxies (Menci et al. 2016a,b), or stellar ages of MW satellites (Chau, Mayer & Governato 2016). As a rule-of-thumb, the current Lyman-α limits are $m_{\rm TH} \gtrsim 3$ keV while all other limits are around $m_{\rm TH} \gtrsim 2$ keV or weaker.[7]

[7] The constraints from Lyman-α rely on assumptions about the temperature evolution of the intergalactic medium. Using very high-redshift quasars and assuming a power-law dependence for the temperature, Viel et al. (2013) obtained the limit $m_{\rm TH} \gtrsim 3.3$ keV at 95 per cent CL. This limit is weakened by about 1 keV if the power-law evolution of the temperature is replaced

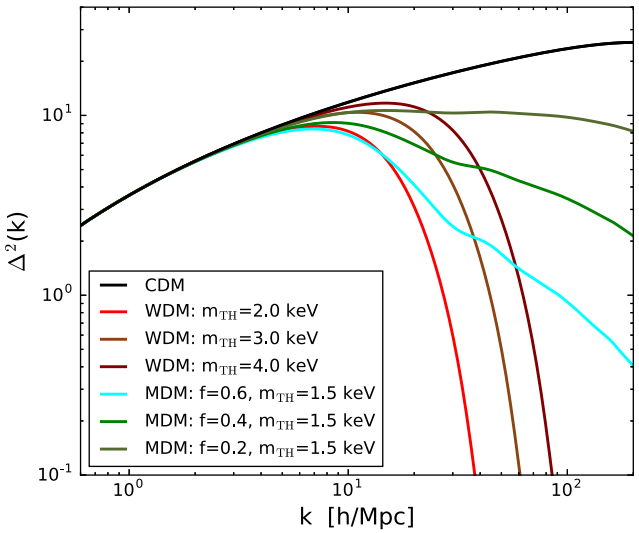

Figure 2. Linear dimensionless power spectra $\Delta^2(k) \equiv k^3 P_{\rm lin}(k)/(2\pi^2)$ of the CDM, WDM, and MDM models discussed in this paper.

In this paper, we study the representative cases of $m_{\rm TH} = 2$, 3, and 4 keV, where the first is in tension with the Lyman-α data but consistent with other limits and the latter two are, roughly speaking, in agreement with observations. The linear power spectra of these models are plotted in Fig. 2 for illustration. They are indistinguishable from CDM at large scales (low wavenumber k) but become strongly suppressed towards smaller scales (high k). The suppression scale only depends on the thermal mass ($m_{\rm TH}$) of the WDM model.

4.1 Halo profiles

It has been shown in the past that WDM models with realistic particle masses do not produce halo cores large enough to be observable (Kuzio de Naray et al. 2010; Villaescusa-Navarro & Dalal 2011; Maccio et al. 2012; Shao et al. 2013). Instead, the haloes are well described by NFW profiles with a modified concentration–mass relation (Schneider 2015; Bose et al. 2016; Ludlow et al. 2016). Rather than monotonically rising towards small masses, as is the case for CDM, the WDM concentrations turn over and decrease again with a maximum around dwarf galaxy scales, with the exact position depending on the DM particle mass (Eke, Navarro & Steinmetz 2001; Schneider et al. 2012).

The modified concentrations of the WDM model affect the calculation of the maximum circular velocities. In the left-hand panel of Fig. 3, we show the WDM (with $m_{\rm TH} = 3$ keV) circular velocity profiles (brown lines) fitted to the data points of the outermost rotation measurement from the selected sample galaxies (black symbols). The lower concentrations at small scales lead to larger values of $v_{\rm max}$ further out in radius compared to the case of CDM (black lines). This means that in WDM models, a galaxy with a given $v_{\rm rot}$ can be fit into a more massive DM halo compared to CDM.

by an abrupt jump in temperature around $z \sim 5$ (Viel et al. 2013; Garzilli, Boyarsky & Ruchayskiy 2015). An even stronger constraint of $m_{\rm TH} \gtrsim 4.1$ keV (at 95 per cent CL) has been found by Baur et al. (2016) using high-redshift quasars from the BOSS survey. This tight limit is, however, relaxed to $m_{\rm TH} \gtrsim 3.0$ keV if cosmological parameters from *Planck* are assumed (instead of the internal parameters from BOSS).

In the middle panel of Fig. 3, we illustrate the difference between WDM and CDM in terms of the relation between $v_{\rm max}$ and $v_{\rm rot}$ (brown and black symbols), where the error bars indicate the sensitivity of $v_{\rm max}$ to variations in concentration.[8] For WDM, the data are better fitted by a quadratic fit (brown line) compared to a linear fit for CDM (black line).

The right-hand panel of Fig. 3 shows the $v_{\rm max}$–$M_{\rm bar}$ relation of the selected galaxy sample for both WDM and CDM (full and empty triangles). While the CDM relation is well described by a linear fit (black solid line), the data bends downwards at small velocities for WDM. We capture this downturn by applying the average shift of $v_{\rm max}$ between WDM and CDM (i.e. the vertical separation between black and brown lines in the middle panel) to the linear fit from CDM, yielding $\mathcal{M}(v_{\rm max})$ for WDM.

Next, we perform a likelihood ratio analysis to find the model $\mathcal{M}_{\rm supp}(v_{\rm max})$ with the maximum baryon suppression. As for CDM, this corresponds to the model with the largest value of $v_{\rm s}$ allowed by the data at the 3σ CL. This model is shown as dashed brown line in Fig. 3. The brown cross indicates the critical velocity ($v_{\rm c}$) where the line of maximal suppression crosses the completeness limit of the K13 sample. Below this scale, the observed abundance of galaxies could be reduced due to baryonic processes (see equation 12).

In Fig. 3, we only illustrate the case of WDM with $m_{\rm TH} = 3$ keV for brevity. Note, however, that other WDM models show very similar trends with increasing discrepancies between WDM and CDM for decreasing thermal-relic mass $m_{\rm TH}$.

4.2 Velocity function

Assuming a WDM scenario affects the VF in several non-trivial ways. First of all, the predicted VF is flatter in WDM than in CDM due to a combination of lower halo abundance and lower concentrations. The former reduces the number of observable galaxies while the latter lowers the maximum circular velocity at a given mass scale. Second, the $v_{\rm max}$-corrected VF from observations becomes steeper in WDM compared to CDM, the reason being the modified relation between $v_{\rm rot}$ and $v_{\rm max}$. This is again a direct consequence of the reduced concentrations which cause galaxies of a given $v_{\rm rot}$ to inhabit larger haloes. Both effects are expected to improve the agreement between theory and observations for WDM compared to CDM.

In Fig. 4, we show the VF of three different WDM models with thermal masses of $m_{\rm TH} = 2$ keV (top left), $m_{\rm TH} = 3$ keV (top right), and $m_{\rm TH} = 4$ keV (bottom left) as well as the VF for CDM (bottom right). Both the flattening of the predicted and the steepening of the observed $v_{\rm max}$–VF are well visible in the plot, the effects becoming more pronounced for decreasing values of $m_{\rm TH}$.

In each panel of Fig. 4, the predicted VF with no baryon effects and with maximum baryon effects are shown as solid and dashed lines. While *baryon depletion* induces a horizontal shift towards small velocities, *baryon suppression* leads to a turnover of the VF below a characteristic velocity ($v_{\rm c}$). The value of $v_{\rm c}$ depends on the model and becomes larger for a smaller thermal relic mass ($m_{\rm TH}$). In general, the area between these lines (hatched area) illustrates the theoretical uncertainty related to unknown baryon effects.

It is obvious from Fig. 4 that the WDM models lead to a better match between theory and observations than CDM. For the models

[8] The sizes of the error bars show the maximum variation in $v_{\rm max}$ if the concentrations are raised or lowered by 1σ with respect to the mean value. They do not include observational uncertainties.

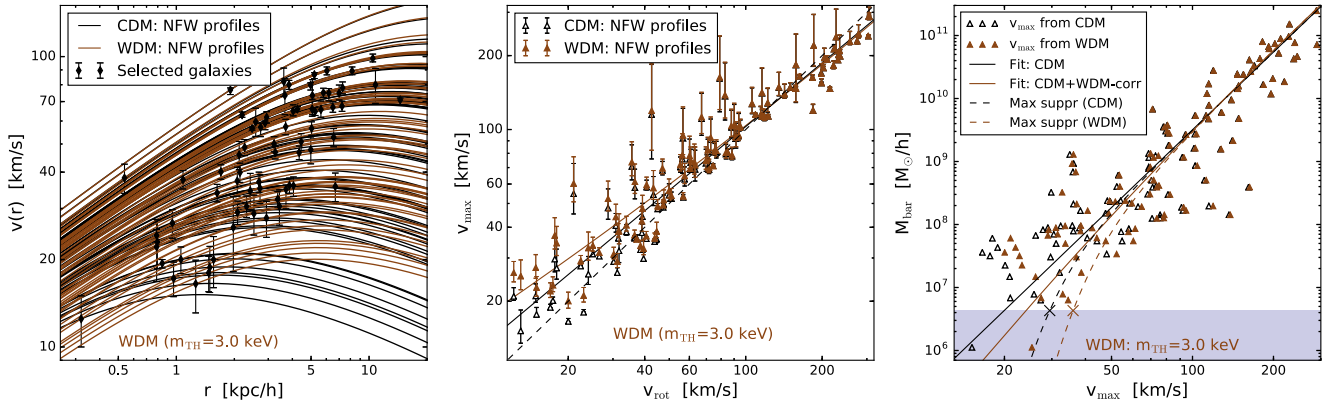

Figure 3. *Left:* NFW velocity profiles (assuming average concentrations) for CDM (black lines) and WDM (brown lines) fitted to the observed circular velocities of the selected galaxy sample (black symbols with error bars). *Middle:* relation between v_{rot} and v_{max} based on the fits from the left-hand panel (empty triangle for CDM and full triangles for WDM). The error bars indicate the change of v_{max} when the halo concentrations are varied by 1σ around the mean. The black and brown lines are linear and quadratic least-squares fit, respectively. The rotational velocity (v_{rot}) is defined via H I line width (w_{50}) and galaxy inclination (i), i.e. $v_{rot} \equiv w_{50}/(2 \sin i)$. *Right:* BTF relation based on v_{max} from CDM and WDM (empty and full triangles). The solid black line is a linear fit to the BTF relation for CDM. The brown line is obtained from the black line by replacing the mean values of v_{max} from CDM with the ones from WDM (see lines in middle panel). The dashed lines show the strongest allowed suppression of the BTF (3σ excluded by the data), while the crosses indicate the velocities where these lines cross the K13 survey limit (grey band). The zero baryon suppression (solid lines) and the maximum baryon suppression (dashed lines) represent the allowable range of galaxy formation models that we will later use in calculating the VF.

with $m_{TH} = 2$ and 3 keV, the v_{max}–VF from observations overlaps with the theory prediction (given the uncertainties in baryonic effects). For the coolest model with $m_{TH} = 4$ keV, a small tension between observation and theory starts to be visible around $v_{max} \sim 40\,km\,s^{-1}$, but the discrepancy is still significantly smaller than for the case of CDM (reproduced in the bottom-right panel of Fig. 4).

It is remarkable that WDM models with thermal masses of $m \gtrsim 3$ keV are able to solve (or at least significantly alleviate) the problem of the VF. These models are cold enough to agree with the very stringent bounds from the Lyman-α forest, therefore offering a truly viable alternative to standard CDM. Former studies have argued that only extreme WDM scenarios, which are in conflict with constraints from Lyman-α and MW satellite counts, are able to solve the mismatch of the VF (Zavala et al. 2009; Papastergis et al. 2011; Schneider et al. 2014; Klypin et al. 2015; Papastergis et al. 2015). These studies, however, did not account for suppression effects from the baryon sector nor for the correction of the observed VF due to larger values of v_{max} for WDM.

4.3 The special case of sterile neutrinos

The sterile (or right-handed) neutrino is often considered as the prime candidate for WDM. It is a well-motivated hypothetical particle based on a straightforward extension of the standard model neutrino sector. Sterile neutrinos can only play the role of DM if their mass is in the keV range, otherwise they would either not cluster enough or decay too quickly (see e.g. Adhikari et al. 2017).

A popular way to produce sterile neutrino DM in the early universe is via resonant mixing with active neutrinos (Shi & Fuller 1999; Abazajian, Fuller & Patel 2001; Asaka, Blanchet & Shaposhnikov 2005). This production mechanism differs from thermal freeze-out and does not lead to Fermi–Dirac-like momentum distributions. As a result, the suppression in the power spectrum can be somewhat shallower than for the case of the standard (thermal relic) WDM, at least for parts of the parameter space (Ghiglieri & Laine 2015; Venumadhav et al. 2016).

A similar effect is observed if sterile neutrino DM is produced via the decay of heavy scalar singlets (Kusenko 2006; Shaposhnikov & Tkachev 2006). Depending on the coupling of the scalar to the standard model and on the decay width, the resulting sterile neutrino momentum distribution can strongly differ from a Fermi–Dirac function and may lead to shallower suppressions of the power spectra (Merle & Totzauer 2015; Merle, Schneider & Totzauer 2016; Konig, Merle & Totzauer 2016).

In terms of the VF, sterile neutrino DM is expected to show qualitatively similar effects to the thermal-like WDM models (see e.g. Lovell et al. 2017). At the quantitative level, some differences are expected due to changes in the shape of the power spectra (Schneider 2016). A detailed investigation of the effects of sterile neutrino DM on the VF will be performed in future work.

5 MIXED DARK MATTER

A straightforward extension to the dark sector is to assume more than one DM species. There is a wealth of possibilities for MDM scenarios including different particle species with different kinds of couplings. These range from two unrelated purely gravitationally interacting DM species to phenomenologically rich scenarios which mirror the baryonic sector. In fact, one could argue that we are already confronted with an MDM universe, since neutrinos are massive, behave exactly like a DM fluid, and have a non-negligible effect on structure formation.

In this section, we limit ourselves to the simple case of a mixture between CDM and (thermal relic) WDM (see e.g. Borgani, Masiero & Yamaguchi 1996; Palazzo et al. 2007; Boyarsky et al. 2009). This is a hypothetical model with the advantage of yielding a large variety of suppressed power spectra. Similar to WDM, it reduces the number of small galaxies (Anderhalden et al. 2013; Maccio et al. 2013) and produces halo profiles with smaller concentrations (Schneider 2015), but the affected mass range can be much larger.

In addition to the particle mass of the warm component (m_{TH}), the MDM model is characterized by the mass fraction of warm to cold species, i.e. $f = \Omega_{WDM}/(\Omega_{WDM} + \Omega_{CDM})$. In this paper, we investigate three cases with the same WDM particle mass of

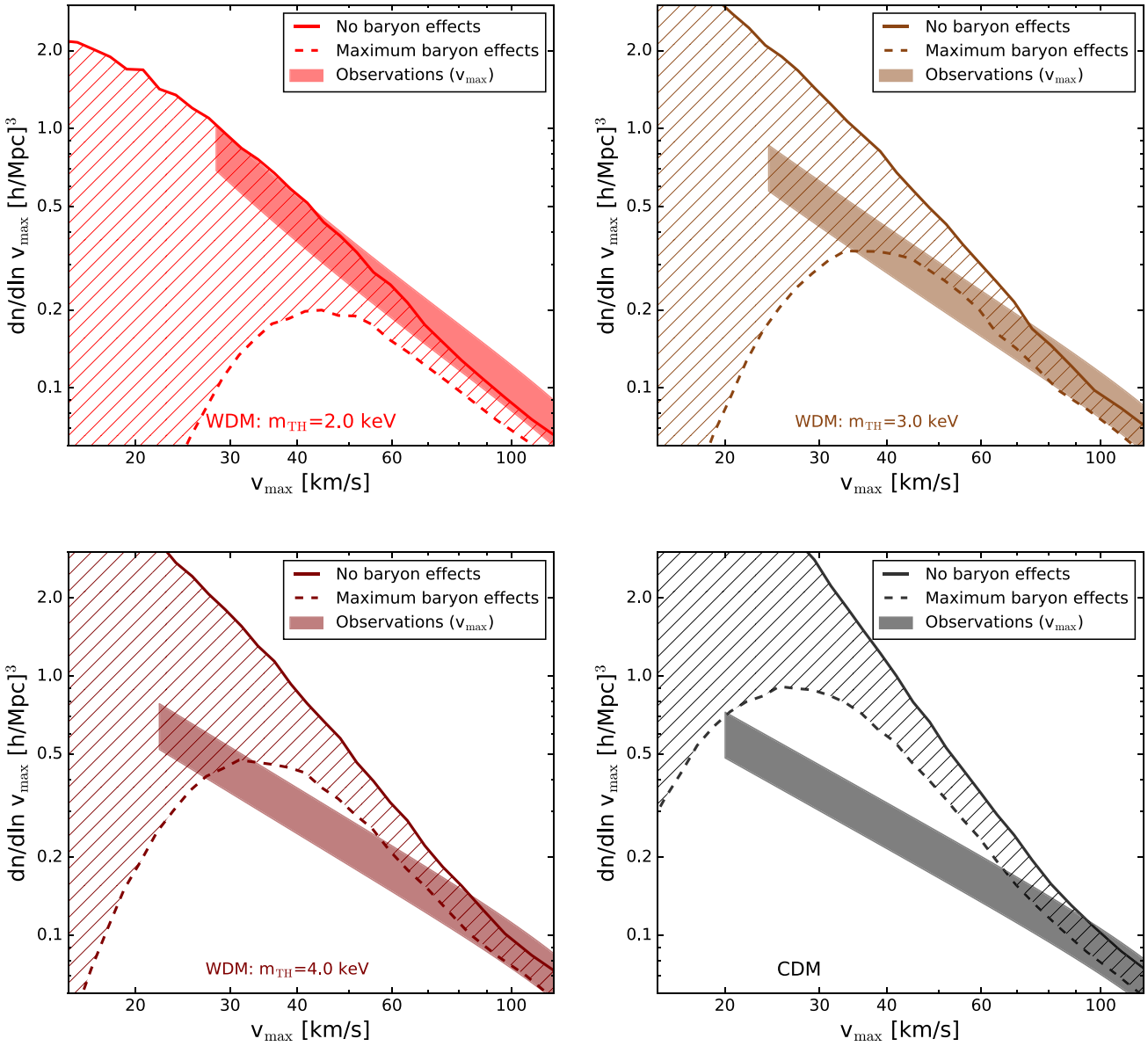

Figure 4. VFs for (thermal relic) WDM models with $m_{\rm TH} = 2$ keV (top left), $m_{\rm TH} = 3$ keV (top right), and $m_{\rm TH} = 4$ keV (bottom left). The CDM model is shown on the bottom right for comparison. The hatched area between solid and dashed lines illustrates the theoretical uncertainty related to unknown baryon processes (see Section 3.2). The agreement between observations (shaded band, corrected for $v_{\rm max}$) and theory is significantly better for WDM than for CDM. Given the uncertainties, WDM models between $m_{\rm TH} \sim 1.8$ and 3.5 keV are able to match the observed abundance of galaxies.

$m_{\rm TH} = 1.5$ keV and different fractions $f = 0.2$, 0.4, and 0.6. Of course, this only covers a very small fraction of the full MDM parameter space, but it serves as an illustration for the kind of corrections that can be expected for other combinations of MDM.

The linear power spectra of the three MDM models are shown in Fig. 2. They are suppressed with respect to CDM but the shape of the suppression is much shallower than for WDM, spanning many orders of magnitudes in length-scale.

5.1 Halo profiles

The haloes of the MDM scenario are well described by NFW profiles (Anderhalden et al. 2012; Maccio et al. 2013) with reduced concentrations at small scales. The shape of the concentration–mass relation can again be directly obtained from the linear power spectrum of MDM by assigning the same concentrations to haloes with the same collapse redshift (see Schneider 2015).

Fig. 5 shows the velocity profiles (left-hand panel) as well as the $v_{\rm max}$–$v_{\rm rot}$ dependence (middle panel) for a MDM model with $f = 0.4$ and $m_{\rm TH} = 1.5$ keV. Similarly to the case of WDM, the values of $v_{\rm max}$ are increased for small galaxies in MDM with respect to CDM. The resulting relation between $v_{\rm max}$ and $v_{\rm rot}$ is well fitted by a curved line which flattens out towards small velocities. The flattening starts at slightly larger scales than for WDM, owing to the smaller mass of the warm component.

The maximum allowed baryon suppression for MDM is obtained in the same way as for the WDM scenario (see Section 4.1). First, we define the function $\mathcal{M}$ describing the $v_{\rm max}$–$M_{\rm bar}$ relation of galaxies in MDM. Then, we determine the maximum allowed baryon

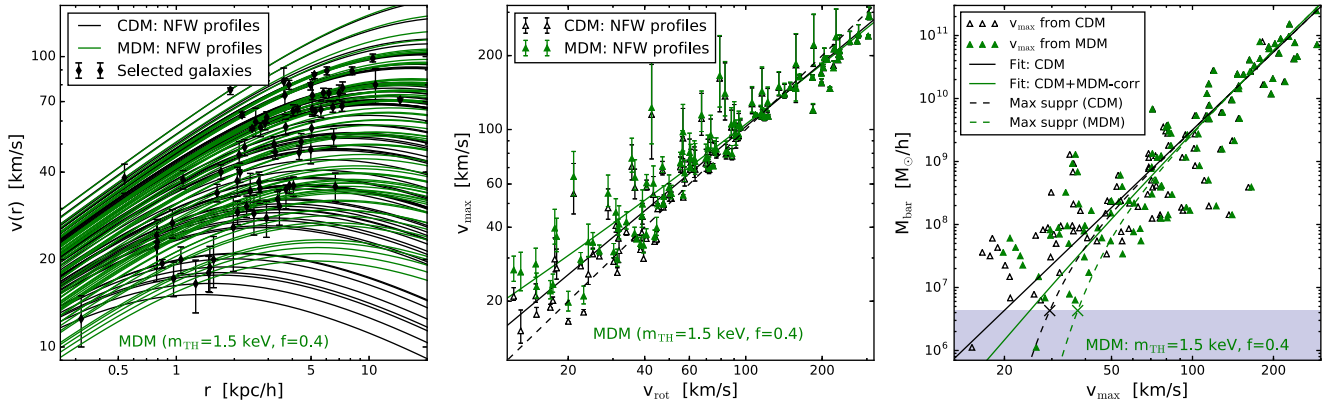

Figure 5. *Left:* NFW velocity profiles (assuming average concentrations) for CDM (black lines) and MDM (green lines) fitted to the observed circular velocities of the selected galaxy sample (black symbols with error-bars). *Middle:* relation between $v_{\rm max}$ and $v_{\rm rot}$ based on the fits from the left-hand panel (empty triangle for CDM, full triangles for MDM). The error bars indicate the change of $v_{\rm max}$ when the halo concentrations are varied by 1σ around the mean. The black and green lines are linear and quadratic least-squares fit, respectively. The rotational velocity ($v_{\rm rot}$) is defined via H I line width (w_{50}) and galaxy inclination (i), i.e. $v_{\rm rot} \equiv w_{50}/(2\sin i)$. *Right:* BTF relation based on $v_{\rm max}$ from CDM and MDM (empty and full triangles). The solid black line is a linear fit to the BTF relation for CDM. The green line is obtained from the black line by replacing the mean values of $v_{\rm max}$ from CDM with the ones from MDM (see lines in middle panel). The dashed lines show the strongest allowed suppression of the BTF (3σ excluded by the data), while the crosses indicate the velocities where these lines cross the K13 survey limit (grey band). The zero baryon suppression (solid lines) and the maximum baryon suppression (dashed lines) represent the allowable range of galaxy formation models that we will later use in calculating the VF.

suppression using the likelihood ratio analysis. The corresponding function ($\mathcal{M}_{\rm supp}$), is shown as dashed green line in Fig. 5. Finally, we use the completeness limit of the K13 sample to determine the characteristic velocity $v_{\rm c}$ (green cross). The value of $v_{\rm c}$ is model dependent (increasing with higher fractions f) and sets the largest scale at which the VF could be affected by baryon suppression.

For brevity, we illustrate only one MDM scenario in Fig. 5. However, other models show similar trends with growing discrepancies between MDM and CDM for increasing fraction f or decreasing mass $m_{\rm TH}$.

5.2 Velocity function

Within the MDM scenario, the VF is affected in a similar way to the case of WDM. First of all, the predicted VF is shallower than in the case of CDM, owing to a combination of reduced halo abundance and concentrations. Second, the observed and $v_{\rm max}$-corrected VF becomes steeper because of higher estimates of $v_{\rm max}$ in MDM as opposed to CDM.

Both effects are visible in Fig. 6, where we plot the MDM models with $m_{\rm TH} = 1.5$ keV and $f = 0.6$ (top left), $f = 0.4$ (top right), and $f = 0.2$ (bottom right). Not surprisingly, all three models provide a much better match between theory and observations than in the case of CDM (bottom right). For the first two models, there is full agreement between ($v_{\rm max}$-corrected) observations (shaded bands) and theory predictions that include the uncertainties of baryon effects (hatched areas, bracketed by the solid and dashed lines). A small tension starts to be visible for the model with $f = 0.2$, but the discrepancy between theory and observations remains significantly smaller than for the case of CDM.

The example of MDM illustrates that many ADM models have the potential to alleviate the problem of the overabundance of field galaxies, provided they suppress perturbations at small scales. In Section 7, we will briefly discuss some other models with similar characteristics.

6 SELF-INTERACTING DARK MATTER

The concept of SIDM became popular after Spergel & Steinhardt (2000) showed that it could provide a better match to dwarf galaxy observations than the standard CDM model. However, it was soon realized that strong self-interactions are in conflict with observations at the scale of galaxy clusters, thereby ruling out the most simple SIDM scenarios (Yoshida et al. 2000; Miralda-Escude 2002). More recently, the SIDM model regained popularity thanks to the realization that previous limits were set too stringently (Rocha et al. 2013; Peter et al. 2013), and that velocity-dependent SIDM models easily evade limits from clusters while being well motivated by particle physics (Feng et al. 2009; Feng, Kaplinghat & Yu 2010; Loeb & Weiner 2011). In addition, some observational studies based on strong lensing found offsets between the mass centres of the stellar and the DM components in clusters, which could be explained by SIDM models (Williams & Saha 2011; Massey et al. 2015).

Concerning the VF, only SIDM with velocity-dependent cross-sections has the potential to reduce the discrepancy on small scales without modifying the large scales. All velocity-independent models alter both small and large scales and can therefore be discarded as a solution to the observed discrepancy of the VF. For the velocity-dependent cross-section, we follow (Feng et al. 2010; Loeb & Weiner 2011) and assume a Yukawa force interaction leading to a scattering cross-section

$$\frac{\sigma}{\sigma_{\rm m}} \simeq \begin{cases} \frac{4\pi}{22.7}\beta^2 \ln(1+\beta^{-1}) & \beta < 0.1 \\ \frac{8\pi}{22.7}\beta^2 (1+1.5\beta^{1.65})^{-1} & 0.1 < \beta < 10^3 \\ \frac{\pi}{22.7}(\ln\beta + 1 - \frac{1}{2}\ln^{-1}\beta)^2 & \beta > 10^3 \end{cases} \quad (13)$$

where m is the mass of the force carrier and $\beta \equiv \pi v_{\rm m}^2/v^2$. The SIDM model has two free parameters given by ($\sigma_{\rm m}/m$) and $v_{\rm m}$.

In Fig. 7, we show how the SIDM cross-sections depend on velocity for the models studied in this paper. All cross-sections are largest at low particle velocities and strongly reduced at large velocities, showing their potential to simultaneously produce significant cores for dwarf galaxies while evading galaxy cluster constraints.

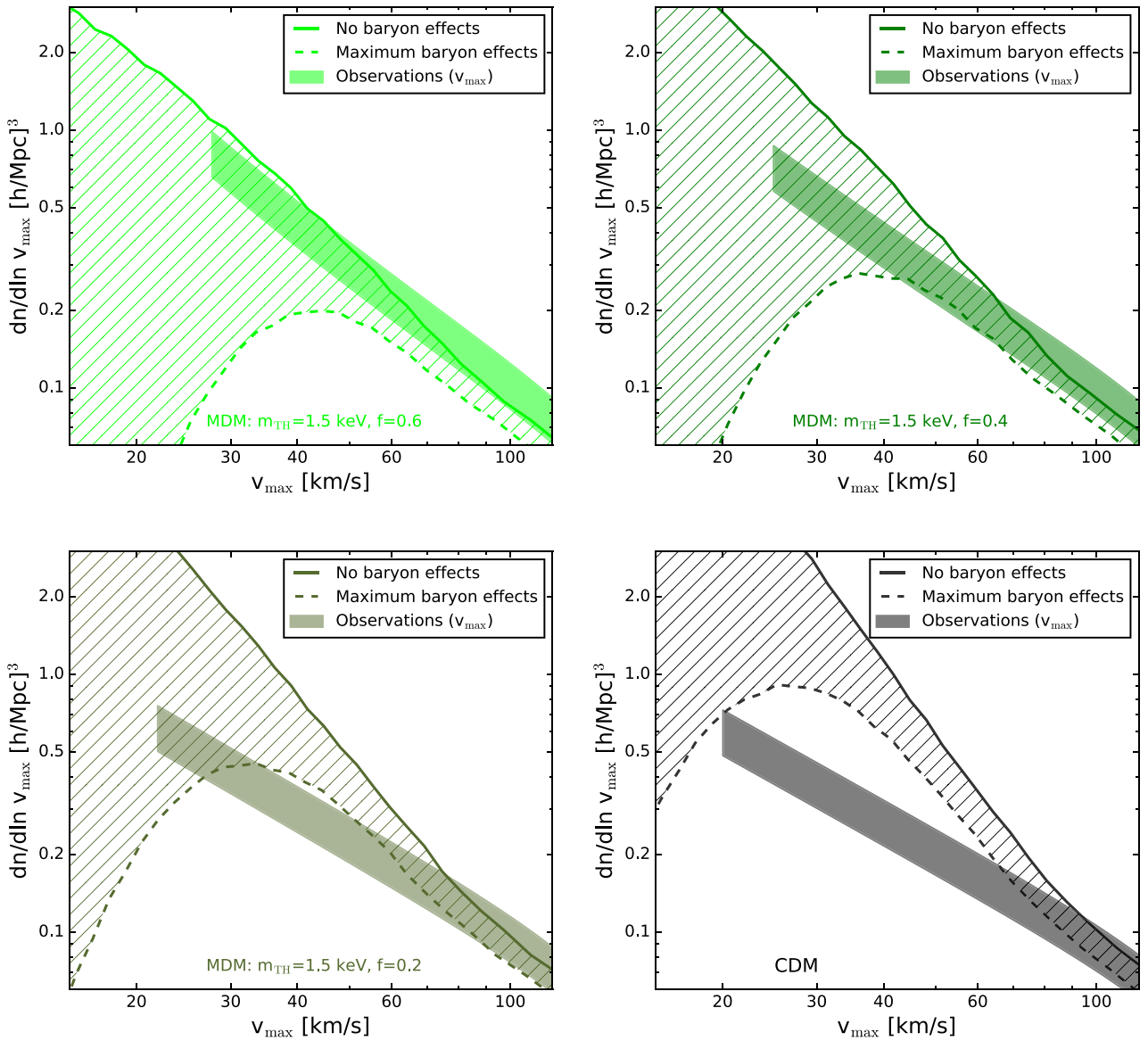

Figure 6. VFs for MDM models (i.e. a mixture of WDM and CDM species) with $m_{\rm TH} = 1.5$ keV and a varying WDM fraction of $f = 0.6$ (top left), $f = 0.4$ (top right), and $f = 0.2$ (bottom left). The CDM model is shown on the bottom right for comparison. The hatched area between solid and dashed lines illustrates the theoretical uncertainty related to unknown baryon processes (see Section 3.2). The agreement between observations (shaded band, corrected for $v_{\rm max}$) and theory is significantly better for MDM than for CDM.

The light grey band in Fig. 7 shows the region of parameter space where the TBTF problem is potentially alleviated due to the reduced central densities of haloes. The dark grey band indicates the region where the largest halo cores are expected. Above this scale, core-collapse starts to dominate, effectively reducing the core sizes despite even larger cross-sections (see Elbert et al. 2015).

The cross-sections from equation (13) have become the standard prescription for velocity-dependent SIDM in the literature. However, there are other models with shallower velocity dependence that are equally justified from a particle physics point of view (see Kaplinghat, Tulin & Yu 2016).

6.1 From cross-sections to halo profiles

The most striking feature of SIDM models in contrast to the CDM scenario is the flattening of the inner part of DM halo density profiles, which is a result of multiple scattering processes in high-density regions. In previous work, SIDM haloes were described by Burkert profiles (Burkert 2000; Zavala, Vogelsberger & Walker 2013). This profile provides a good fit to the inner parts of a halo, but slightly deviates towards large radii (Rocha et al. 2013). In this paper, we use the R16 profile (Read et al. 2016b) instead, which has the advantage of becoming an NFW profile well beyond the core radius. So far, the R16 profile has only been applied to core transformations induced by baryons, but we show in Appendix B that it also provides a very accurate fit to simulated SIDM haloes from the literature.

To assign concentrations to SIDM haloes, we use the same relation as for CDM, implicitly assuming that the rare collisions, well beyond the core radius, have a negligible effect on the profile. This assumption seems reasonable but should be tested in the future.

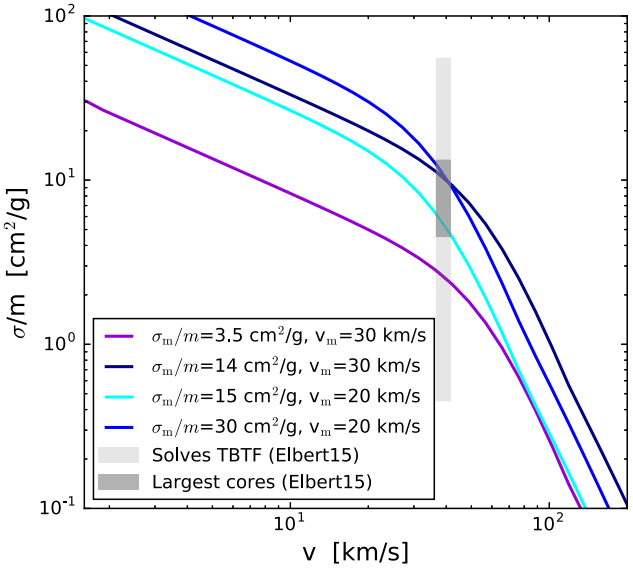

Figure 7. Velocity-dependent cross-sections for the SIDM models studied in this paper. The light and dark grey bands indicate the range of scales where the TBTF problem is alleviated and where the largest cores are expected (see Elbert et al. 2015). The purple model corresponds to a scenario investigated in Vogelsberger, Zavala & Loeb (2012).

There is a direct relation between the cross-section of SIDM and the average core size of haloes. Unfortunately, no simulation-based study has ever investigated this connection systematically. It is, however, possible to determine the approximate core size (r_c) analytically, using estimates of the average DM interaction rate (Γ). Assuming that a fixed number of interactions per *Hubble* time is required to produce a core, we obtain the relation

$$\rho_{\rm nfw}(r_c|M)\frac{\langle\sigma v\rangle(r_c)}{m} \simeq \Gamma. \quad (14)$$

Here, we closely follow the approach of Dooley et al. (2016), but we assume an NFW profile (instead of a Hernquist profile) and we use an interaction rate of $\Gamma = 0.4\,{\rm Gyr}^{-1}$ (instead of $\Gamma = 1.0\,{\rm Gyr}^{-1}$). The latter gives a better match to simulated SIDM haloes in combination with the R16 core definition (see Appendix B for more details). Following Vogelsberger et al. (2012), the average velocity-weighted cross-sections are given by the integral over the Maxwell–Boltzmann distribution, i.e.

$$\langle\sigma v\rangle(r) = \frac{1}{\sqrt{4\pi}\sigma^3_{\rm vel}(r)}\int {\rm d}v v^2(\sigma v)\exp\left[-\frac{v^2}{4\sigma^2_{\rm vel}(r)}\right]. \quad (15)$$

The velocity dispersion $\sigma_{\rm vel}$ depends on the halo profile and can be calculated by solving the isotropic Jeans equation, ${\rm d}(\rho\sigma^2_{\rm vel})/{\rm d}r = -\rho{\rm d}\phi/{\rm d}r$ (where ϕ is the potential and ρ is the DM density), leading to

$$\sigma^2_{\rm vel}(r) = \frac{G}{\rho_{\rm nfw}(r)}\int_r^\infty {\rm d}x\frac{M_{\rm nfw}(x)\rho_{\rm nfw}(x)}{x^2}. \quad (16)$$

equation (14) can be combined with equations (15) and (16) to find a relation between core size (r_c) and halo mass (M). Since this relation is calibrated to SIDM simulations from the literature (via the interaction rate parameter Γ), it is expected to provide reasonably accurate results over a large range of scales. A similar approach to estimate halo cores from SIDM cross-sections can be found in Kaplinghat et al. (2016).

In the top-left panel of Fig. 9, we plot the relation between core size (r_c) and halo mass (M) for the SIDM models studied in this paper. The halo cores only vary by a factor of a few over a large range of mass scales. This is in strong contrast to SIDM models with velocity-independent cross-sections, where there is a strong power-law dependence with core size increasing with halo mass (see Dooley et al. 2016).

Given a halo density profile (i.e. equation 1) and a core radius (r_c) for SIDM, we can perform profile fits to all the galaxies in the selected sample. In the left-hand panel of Fig. 8, we plot the velocity profiles of CDM (grey lines) and SIDM (with $\sigma_m/m = 14\,{\rm cm}^2\,{\rm g}^{-1}$, $v_m = 30\,{\rm km\,s}^{-1}$, dark-blue lines) fitted to the observed circular velocities $v_{\rm out}$ at $r_{\rm out}$ (symbols). The large SIDM cores lead to steep velocity profiles at small radii, requiring that small galaxies inhabit larger haloes. This becomes even more evident in the middle panel of Fig. 8 where the maximum circular velocity ($v_{\rm max}$) is plotted against the observed rotation velocity derived from the H I line width ($v_{\rm rot}$). There is a strong flattening and an increase of scatter visible in $v_{\rm max}$–$v_{\rm rot}$ relation towards small velocity scales (fitted by the solid dark-blue line). The latter is in line with the recently predicted higher variability of SIDM rotation curves with respect to CDM (Elbert et al. 2016; Kamada et al. 2016; Creasey et al. 2016)[9]

Finally, we plot the $v_{\rm max}$–$M_{\rm bar}$ relation for SIDM in the right-hand panel of Fig. 8. The relation is identical to the one of CDM at large velocity scales and shows both a downturn and an increase of scatter towards smaller scales. Following the approach used for the WDM and MDM models, we define the function $\mathcal{M}(v_{\rm max})$ by modifying the linear fit from CDM (using the average shift between $v_{\rm max}$ from SIDM and CDM, i.e. the vertical separation between the black and dark-blue lines in the middle panel). This leads to the solid blue line in the right-hand panel of Fig. 8. The model with maximum baryon suppression, $\mathcal{M}_{\rm supp}(v_{\rm max})$, is shown as a dashed blue line. The line crosses the completeness limit of the K13 sample at the characteristic velocity, v_c (dark-blue cross). The value of v_c is model dependent (growing for increasing cross-sections) and sets the maximum scale at which the theoretical abundance of galaxies hosted by SIDM haloes could be affected by suppression effects from baryonic processes.

Note that only one particular SIDM model is illustrated in Fig. 8 for brevity. Other scenarios show similar trends with growing differences between SIDM and CDM for larger particle cross-sections.

6.2 Velocity function

In contrast to WDM and MDM, the SIDM model does not lead to a reduction of the halo abundance but simply modifies their inner structure. Obviously, the disagreement between the observed and predicted VF of galaxies can only be solved with substantially larger cores than the ones induced by baryons, which we showed in Section 2.2 to be not large enough to affect the result.

In Fig. 9, we plot the VF of the four SIDM scenarios introduced above. The two models with large cross-sections and cores above $r_c \sim 2$ kpc are able to fully reconcile theory with observations (see top panels). The third model is also marginally consistent with observations while the last model only slightly reduces the tension with respect to CDM.

[9] The larger scatter of SIDM rotation curves is the result of an interplay between the stellar and the DM components. Depending on the number and distribution of stars in the halo centres, the DM component may or may not experience core collapse, resulting in a large diversity of profiles. In principle, this effect is testable by combining H I rotation curves with the observed stellar density profiles.

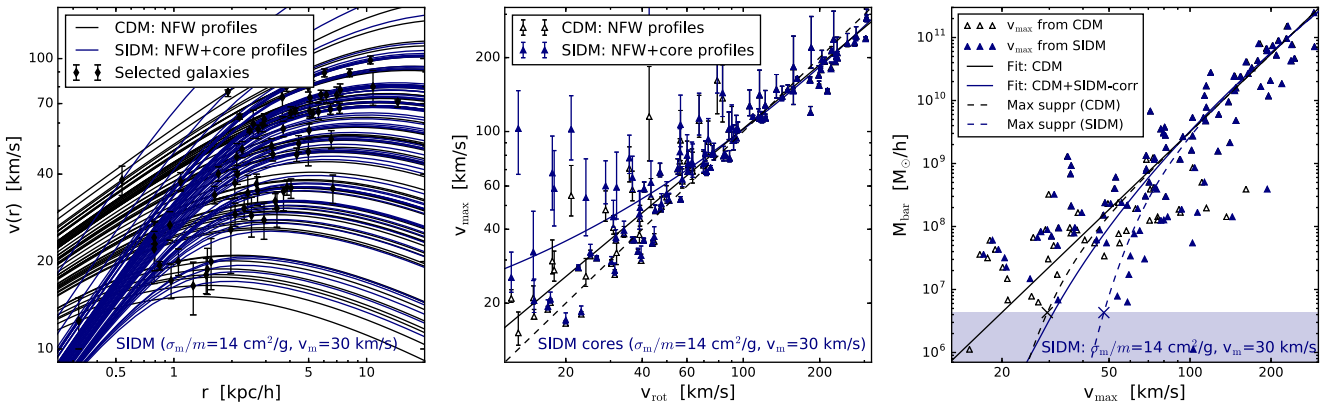

Figure 8. *Left:* NFW velocity profiles (assuming average concentrations) for CDM (black lines) and SIDM (dark-blue lines) fitted to the observed circular velocities of the selected galaxy sample (black symbols with error-bars). *Middle:* relation between $v_{\rm max}$ and $v_{\rm rot}$ based on the fits from the left-hand panel (empty triangle for CDM, full triangles for SIDM). The error bars indicate the change of $v_{\rm max}$ when the halo concentrations are varied by 1σ around the mean. The black and brown lines are linear and quadratic least-squares fit, respectively. The rotational velocity ($v_{\rm rot}$) is defined via H I line width (w_{50}) and galaxy inclination (i), i.e. $v_{\rm rot} \equiv w_{50}/(2\sin i)$. *Right:* BTF relation based on $v_{\rm max}$ from CDM and SIDM (empty and full triangles). The solid black line is a linear fit to the BTF relation for CDM. The dark-blue line is obtained from the black line by replacing the mean values of $v_{\rm max}$ from CDM with the ones from SIDM (see lines in middle panel). The dashed lines show the strongest allowed suppression of the BTF (3σ excluded by the data), while the crosses indicate the velocities where these lines cross the K13 survey limit (grey band). The zero baryon suppression (solid lines) and the maximum baryon suppression (dashed lines) represent the allowable range of galaxy formation models that we will later use in calculating the VF.

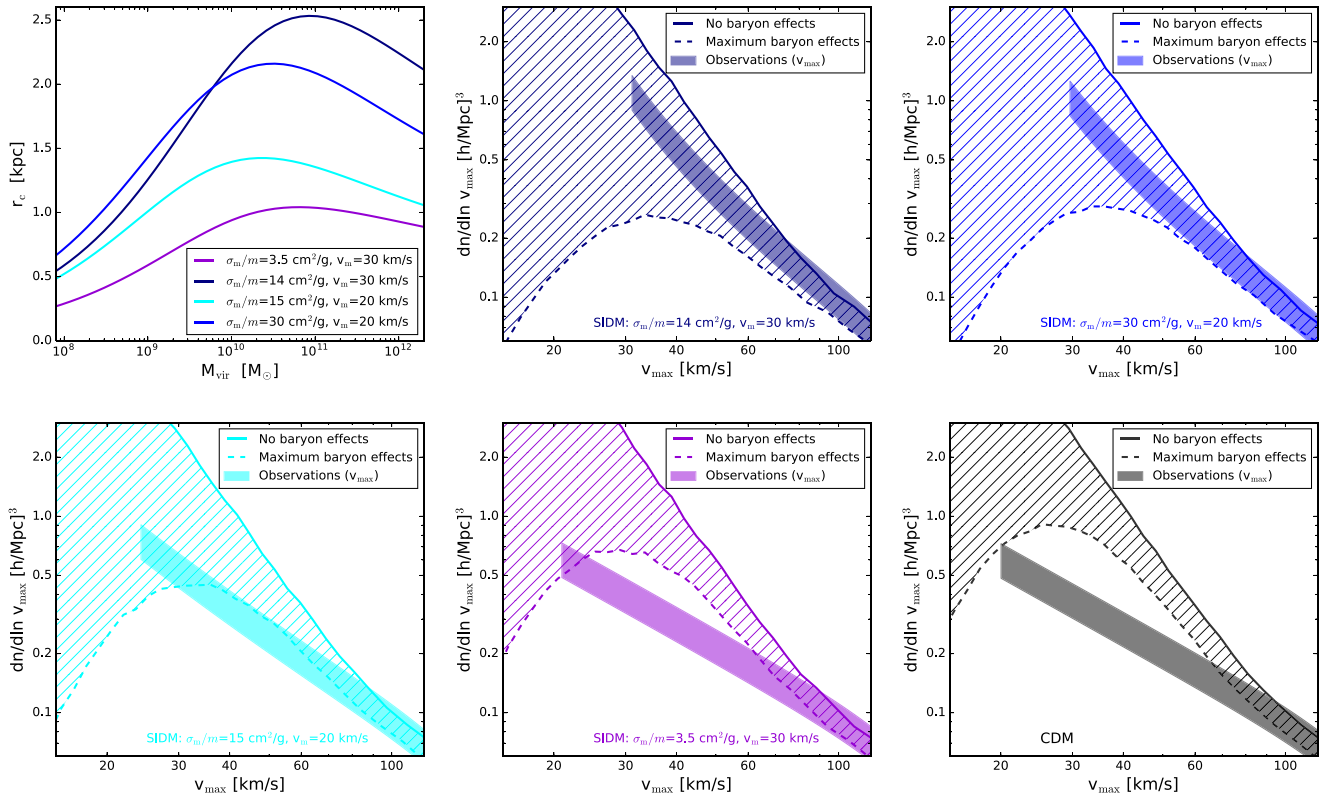

Figure 9. The effective core sizes (top left) and the resulting VFs with respect to $v_{\rm max}$ for SIDM scenarios. The CDM model is shown on the bottom right for comparison. The hatched area between solid and dashed lines illustrate the theoretical uncertainty related to unknown baryonic processes (see Section 3.2). The shaded bands represent the observed VFs corrected for $v_{\rm max}$. Cores above $r_{\rm c} \sim 1-1.5$ kpc are required to solve the mismatch between observations and theory.

In general, the SIDM models lead to a steepening of the observed $v_{\rm max}$–VF, while the predicted VF does not become shallower (as is the case for WDM and MDM). However, the VF can be more strongly suppressed by baryon processes in SIDM as opposed to CDM. This is due to the shallower $v_{\rm max}$–$v_{\rm rot}$ relation and the increased scatter, allowing for a stronger downturn of the $v_{\rm max}$–BTF relation.

While it is possible to solve the 'missing dwarfs' problem with SIDM models, sufficiently strong cross-sections are required, producing cores of $r_{\rm c} \sim 1-1.5$ kpc or larger (at the relevant

dwarf-galaxy scales). It remains to be established whether these models are in agreement with other potential constraints from structure formation. For example, larger cross-sections may lead to faster evaporation of substructures (due to high-velocity encounters with particles from the host halo) which could destroy too many satellites in MW-sized objects. Detailed simulations are required to refute or confirm these concerns.

7 OTHER DARK MATTER CANDIDATES

In addition to the DM models presented in this paper, there are other scenarios with the potential to alleviate the mismatch between the predicted and observed VF. We will now mention some of them, however without providing a quantitative analysis.

An obvious DM candidate with the potential to solve the discrepancy are axion-like particles (ALPs). At very low mass scales of $m \sim 10^{-23}$ eV, ALPs start to form coherent waves of astrophysical length-scales, leading to a suppression of the power spectrum (Hu, Barkana & Gruzinov 2000; Marsh & Silk 2014), and to the formation of soliton cores (Schive et al. 2014). For this reason, ALPs are sometimes referred to as *fuzzy* or *wave* DM. Whether the soliton cores are sufficiently large to be of relevance for the VF is, however, still unclear (Schive et al. 2016; Hui et al. 2017).

Another interesting scenario consists of DM particles coupled to some relativistic fluid like photons, neutrinos, or dark radiation (see e.g. Cyr-Racine & Sigurdson 2013; Boehm et al. 2014; Bringmann et al. 2016). Such a coupling generates a suppression of the power spectrum usually combined with acoustic oscillations at the suppression scale. In terms of effects on the galaxy abundance, a similar behaviour than for WDM or MDM can be expected. It would furthermore be interesting to establish if the dark acoustic oscillation could still be visible in the galaxy VF (see also Buckley et al. 2014).

An interesting new framework to systematically study interactions of the dark sector has recently been presented under the name of ETHOS (Effective THeory of Structure formation, Cyr-Racine et al. 2016; Vogelsberger et al. 2016). Some of the example cases investigated by the authors show models with both suppressed power spectra and significant halo cores. The VF of the local volume could provide an ideal testbed to further study such scenarios.

8 CONCLUSIONS

In the last few years, it became evident that the observed *H I velocity width function* of galaxies in the local volume is in tension with predictions from gravity-only simulations based on the standard model of ΛCDM (e.g. Zavala et al. 2009; Papastergis et al. 2011; Klypin et al. 2015). The discrepancy cannot be fully solved with baryon effects, such as strong UV photoevaporation or supernova feedback, because these processes induce a downturn in the $v_{\rm max}$–$M_{\rm bar}$ relation, which is not observed (Trujillo-Gomez et al. 2016, TG16). This crucial point is further highlighted in Appendix A, where we compare to other work based on AM techniques, simulations, and direct mass estimates from observed dwarf galaxies.

In principle, a number of observational effects could be affecting the kinematic analysis performed in this article, and could therefore alter our conclusions regarding the viability of ΛCDM.[10] For example, the subsample of selected galaxies (with spatially resolved kinematic measurements) used to determine $v_{\rm max}$ could be biased with respect to the full sample used for the VF. This possibility was investigated by TG16, who found no systematic differences between the selected and full samples of local volume galaxies. Another possibility is a bias from inaccurate estimates of galaxy inclinations. Indeed, highly inclined galaxies (which give more accurate measurements of $v_{\rm rot}$) tend to have somewhat smaller velocities at a given baryonic mass. This means, however, that reducing inclination errors would shift the observed VF of $v_{\rm max}$ further away from the CDM predictions, worsening the discrepancy.

A further potential source of error are dwarf galaxies with an extremely low surface brightness falling below the survey detection limit. If such objects exist in large numbers, they could potentially explain the difference between the observed and predicted VF. A detailed discussion of this possibility is given by Klypin et al. (2015). There it is argued that while some of the smallest galaxies with $v_{\rm rot} < 20\,{\rm km\,s^{-1}}$ could potentially stay undetected because of extremely low surface-brightness, they are very unlikely to make up a sizeable fraction of the full population.[11]

Finally, the mismatch between the observed and predicted $v_{\rm max}$–VF could originate from errors in the fitting procedure used to determine $v_{\rm max}$, with the main concern being the possibility of inner DM cores induced by stellar feedback. However, existing estimates produced by baryonic feedback effects (Read et al. 2016b; Di Cintio et al. 2014) have been shown to be too small to produce a significant effect in galaxies with kinematic measurements at large radii (see Fig. 1 and TG16). Finally, the results of this paper depend on the customary assumption that spatially resolved H I rotation measurements (including corrections for turbulence) can be used to probe the halo potential.

A more speculative but intriguing option is that the mismatch between observed and predicted abundance of isolated galaxies points towards an ADM sector. In the present paper, we investigated different DM scenarios such as WDM, MDM, and SIDM, and we showed that they can be in much better agreement with observations. A more detailed summary of the results follows:

(i) The WDM scenario is characterized by a steep cut-off in the initial power spectrum, resulting in two important effects: reduced halo abundance, and lower concentrations. The two effects work together to flatten the predicted $v_{\rm max}$–VF. Furthermore and due to the lower concentrations, observed galaxies are expected to reside in more massive haloes, yielding a steeper relation for the observed $v_{\rm max}$–VF. As a result, observations and theory agree for WDM models with (thermal-relic) masses between $m_{\rm TH} \sim 1.8$ and 3.5 keV (see Fig. 4). This includes lukewarm DM scenarios which are cold enough to avoid the most stringent constraints from the Lyman-α forest. Significantly warmer models with $m_{\rm TH} \lesssim 1.5$ keV are disfavoured by the observed $v_{\rm max}$-VF.

(ii) We also considered a simple MDM scenario with varying fraction of WDM to CDM. This model leads to a wide range of shapes for the power spectrum making MDM an ideal testbed for structure formation. The resulting effect on the VF is qualitatively

[10] For a detailed discussion about potential systematics regarding the galaxy sample, see section 4.1 of Papastergis & Ponomareva (2016).

[11] Furthermore, the most recent searches for extremely low surface brightness dwarfs around massive spirals (including Local Group dwarfs and MW and M31 satellites) find very few objects with $\mu > 27$ mag arcsec^{-2} brighter than $M_{\rm V} = -10$ (see e.g. Merritt, van Dokkum & Abraham 2014). Therefore, most of the dwarfs in the local volume that went undetected due to their low surface brightness should be below the magnitude limit of the catalogue from Karachentsev et al. (2013), and hence would not affect our conclusions.

similar to the case of WDM, except that the flattening can affect a wider range of scales. For large parts of the MDM parameter space, the agreement between observations and theory is highly improved with respect to the case of CDM (see Fig. 6). Again, this includes models that are in agreement with Lyman-α limits.

(iii) The SIDM scenario is qualitatively different from WDM and MDM models in the sense that it does not yield suppressed perturbations at small scales. Instead, the inner parts of halo profiles are flattened due to repeated collisions of DM particles in high-density regions. The flattening of profiles leads to the inevitable conclusion that small galaxies should inhabit more massive haloes compared to CDM. As a result, the observed $v_{\rm max}$–VF becomes steeper for increasing cross-sections. The theoretical abundance of galaxies in SIDM models is unchanged with respect to CDM as long as baryonic processes are neglected. The model including the maximum allowed baryonic suppression, on the other hand, is less constrained than for CDM due to the higher values (and the increased scatter) of $v_{\rm max}$ estimated for observed dwarf galaxies. As a result, SIDM models can fully solve the tension between the predicted and observed $v_{\rm max}$–VF as long as they form sufficiently large DM cores of $r_{\rm c} \gtrsim 1.5$ kpc in dwarf galaxies (see Fig. 9). This is only possible for models with significant cross-sections which have to be velocity-dependent to avoid constraints from galaxy clusters. Whether these models are in agreement with other small-scale observables has yet to be established.

In general, all models that either suppress perturbations at dwarf galaxy scales or flatten the inner DM halo density profiles (or a combination of both) can potentially alleviate the mismatch between predicted and observed abundance of galaxies as a function of $v_{\rm max}$. This includes many more scenarios than the ones studied here. Examples are interacting, decaying, late decoupling, or boson DM. A more detailed investigation of such models is postponed to future work.

Upcoming large area H I surveys, such as the APERTIF survey with the WSRT interferometer and the WALLABY survey with the ASKAP interferometer will provide large samples of dwarf galaxies with spatially resolved velocity information and improved H I sensitivity. This will make it possible to track down the remaining potential systematics related to the profile fitting procedure. The new data will furthermore allow to extend the observed VF to smaller scales, well below 10 km s^{-1} in velocity. This should lead to the discovery of a downturn of both the VF and the $v_{\rm max}$–$M_{\rm bar}$ relation due to the effects of photoevaporation during reionization. Once this effect is known to better accuracy, it will be possible to come up with highly improved constraints for the particle nature of DM.

ACKNOWLEDGEMENTS

We thank Anatoly Klypin for very helpful suggestions on how to improve the present manuscript. AS acknowledges support from the Swiss National Science Foundation (PZ00P2_161363). EP is supported by a postdoctoral fellowship of the Netherlands Research School for Astronomy (NOVA).

REFERENCES

Abazajian K., Fuller G. M., Patel M., 2001, Phys. Rev. D, 64, 023501
Adhikari R. et al., 2017, J. Cosmol. Astropart. Phys., 1701, 025
Anderhalden D., Diemand J., Bertone G., Maccio A. V., Schneider A., 2012, J. Cosmol. Astropart. Phys., 1210, 047
Anderhalden D., Schneider A., Maccio A. V., Diemand J., Bertone G., 2013, J. Cosmol. Astropart. Phys., 1303, 014
Asaka T., Blanchet S., Shaposhnikov M., 2005, Phys. Lett. B, 631, 151
Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, J. Cosmol. Astropart. Phys., 1608, 012
Bekeraitė S. et al., 2016, ApJ, 827, L36
Benson A. J. et al., 2013, MNRAS, 428, 1774
Boehm C., Schewtschenko J. A., Wilkinson R. J., Baugh C. M., Pascoli S., 2014, MNRAS, 445, L31
Borgani S., Masiero A., Yamaguchi M., 1996, Phys. Lett. B, 386, 189
Bose S., Hellwing W. A., Frenk C. S., Jenkins A., Lovell M. R., Helly J. C., Li B., 2016, MNRAS, 455, 318
Boyarsky A., Lesgourgues J., Ruchayskiy O., Viel M., 2009, J. Cosmol. Astropart. Phys., 0905, 012
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40
Bringmann T., Ihle H. T., Kersten J., Walia P., 2016, Phys. Rev. D94, 103529
Brook C. B., Di Cintio A., 2015a, MNRAS, 450, 3920
Brook C. B., Di Cintio A., 2015b, MNRAS, 453, 2133
Brook C., Shankar F., 2016, MNRAS, 455, 3841
Brooks A. M., Papastergis E., Christensen C. R., Governato F., Stilp A., Quinn T. R., Wadsley J., 2017, preprint (arXiv:e-prints) (Brooks17)
Buckley M. R., Zavala J., Cyr-Racine F.-Y., Sigurdson K., Vogelsberger M., 2014, Phys. Rev. D, 90, 043524
Burkert A., 2000, ApJ, 534, L143
Chau A., Mayer L., Governato F., 2016, ApJ, preprint (arXiv:1605.01063)
Creasey P., Sameie O., Sales L. V., Yu H.-B., Vogelsberger M., Zavala J., 2016, MNRAS, 468, 2283
Cyr-Racine F.-Y., Sigurdson K., 2013, Phys. Rev. D, 87, 103515
Cyr-Racine F.-Y., Sigurdson K., Zavala J., Bringmann T., Vogelsberger M., Pfrommer C., 2016, Phys. Rev. D, 93, 123527
de Blok W. J. G., 2010, Adv. Astron., 2010, 789293
Di Cintio A., Brook C. B., Macci A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, MNRAS, 437, 415
Dooley G. A., Peter A. H. G., Vogelsberger M., Zavala J., Frebel A., 2016, MNRAS, 461, 710
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Dutton A. A. et al., 2011, MNRAS, 416, 322
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Elbert O. D., Bullock J. S., Garrison-Kimmel S., Rocha M., Oorbe J., Peter A. H. G., 2015, MNRAS, 453, 29
Elbert O. D., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Graus A. S., Rocha M., 2016, preprint (arXiv:1609.08626)
Fattahi A., Navarro J. F., Sawala T., Frenk C. S., Sales L. V., Oman K., Schaller M., Wang J., 2016, MNRAS, preprint (arXiv:1607.06479)
Feng J. L., Kaplinghat M., Tu H., Yu H.-B., 2009, J. Cosmol. Astropart. Phys., 0907, 004
Feng J. L., Kaplinghat M., Yu H.-B., 2010, Phys. Rev. D, 82, 083525
Fitts A. et al., 2016, MNRAS, preprint (arXiv:1611.02281)
Garzilli A., Boyarsky A., Ruchayskiy O., 2015, preprint (arXiv:1510.07006)
Ghiglieri J., Laine M., 2015, J. High Energy Phys., 11, 171
Gnedin N. Y., 2000, ApJ, 542, 535
Gonzalez A. H., Williams K. A., Bullock J. S., Kolatt T. S., Primack J. R., 2000, ApJ, 528, 145
Governato F. et al., 2012, MNRAS, 422, 1231
Hahn O., Angulo R. E., 2016, MNRAS, 455, 1115
Hobbs A., Read J., Agertz O., Iannuzzi F., Power C., 2016, MNRAS, 458, 468
Horiuchi S., Humphrey P. J., Onorbe J., Abazajian K. N., Kaplinghat M., Garrison-Kimmel S., 2014, Phys. Rev. D, 89, 025017
Hu W., Barkana R., Gruzinov A., 2000, Phys. Rev. Lett., 85, 1158
Hui L., Ostriker J. P., Tremaine S., Witten E., 2017, Phys. Rev. D, 95, 043541
Kamada A., Kaplinghat M., Pace A. B., Yu H.-B., 2016, preprint (arXiv:1611.02716)
Kaplinghat M., Tulin S., Yu H.-B., 2016, Phys. Rev. Lett., 116, 041302
Karachentsev I. D., Makarov D. I., Kaisina E. I., 2013, AJ, 145, 101 (K13)
Karukes E. V., Salucci P., 2017, MNRAS, 465, 4703
Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487
Klypin A. A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Klypin A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Klypin A., Karachentsev I., Makarov D., Nasonova O., 2015, MNRAS, 454, 1798
Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, MNRAS, 457, 4340
Konig J., Merle A., Totzauer M., 2016, J. Cosmol. Astropart. Phys., 1611, 038
Kusenko A., 2006, Phys. Rev. Lett., 97, 241301
Kuzio de Naray R., Martinez G. D., Bullock J. S., Kaplinghat M., 2010, ApJ, 710, L161
Loeb A., Weiner N., 2011, Phys. Rev. Lett., 106, 171302
Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300
Lovell M. R. et al., 2017, MNRAS, 468, 4285
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
Maccio A. V., Paduroiu S., Anderhalden D., Schneider A., Moore B., 2012, MNRAS, 424, 1105
Maccio A. V., Ruchayskiy O., Boyarsky A., Munoz-Cuartas J. C., 2013, MNRAS, 428, 882
Macciò A. V., Udrescu S. M., Dutton A. A., Obreja A., Wang L., Stinson G. R., Kang X., 2016, MNRAS, 463, 69
Marsh D. J. E., Silk J., 2014, MNRAS, 437, 2652
Massey R. et al., 2015, MNRAS, 449, 3393
Menci N., Sanchez N. G., Castellano M., Grazian A., 2016a
Menci N., Grazian A., Castellano M., Sanchez N. G., 2016b, ApJ, 825, L1
Merle A., Schneider A., 2015, Phys. Lett. B, 749, 283
Merle A., Totzauer M., 2015, J. Cosmol. Astropart. Phys., 1506, 011
Merle A., Schneider A., Totzauer M., 2016, J. Cosmol. Astropart. Phys., 1604, 003
Merritt A., van Dokkum P., Abraham R., 2014, ApJ, 787, L37
Miralda-Escude J., 2002, ApJ, 564, 60
Moore B., Ghigna S., Governato F., Lake G., Quinn T. R., Stadel J., Tozzi P., 1999, ApJ, 524, L19
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Obreschkow D., Ma X., Meyer M., Power C., Zwaan M., Staveley-Smith L., Drinkwater M. J., 2013, ApJ, 766, 137
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Onorbe J., Boylan-Kolchin M., Bullock J. S., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., Quataert E., Murray N., 2015, MNRAS, 454, 2092
Pace A. B., 2016, MNRAS, preprint (arXiv:1605.05326)
Palazzo A., Cumberbatch D., Slosar A., Silk J., 2007, Phys. Rev. D, 76, 103511
Papastergis E., Ponomareva A. A., 2016, A&A, 601, A1
Papastergis E., Shankar F., 2016, A&A, 591, A58
Papastergis E., Martin A. M., Giovanelli R., Haynes M. P., 2011, ApJ, 739, 38
Papastergis E., Giovanelli R., Haynes M. P., Shankar F., 2015, A&A, 574, A113
Peter A. H. G., Rocha M., Bullock J. S., Kaplinghat M., 2013, MNRAS, 430, 105
Planck Collaboration XXIV, 2016, A&A, 594, A24
Polisensky E., Ricotti M., 2011, Phys. Rev. D, 83, 043506
Read J. I., Iorio G., Agertz O., Fraternali F., 2016a, MNRAS, 467, 2019 (Read16)
Read J. I., Agertz O., Collins M. L. M., 2016b, MNRAS, 459, 2573
Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Onorbe J., Moustakas L. A., 2013, MNRAS, 430, 81
Sales L. V. et al., 2017, MNRAS, 464, 2419
Sawala T. et al., 2016, MNRAS, 457, 1931
Schive H.-Y., Liao M.-H., Woo T.-P., Wong S.-K., Chiueh T., Broadhurst T., Hwang W. Y. P., 2014, Phys. Rev. Lett., 113, 261302
Schive H.-Y., Chiueh T., Broadhurst T., Huang K.-W., 2016, ApJ, 818, 89
Schneider A., 2015, MNRAS, 451, 3117
Schneider A., 2016, J. Cosmol. Astropart. Phys., 1604, 059
Schneider A., Smith R. E., Maccio A. V., Moore B., 2012, MNRAS, 424, 684
Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573
Schneider A., Anderhalden D., Maccio A., Diemand J., 2014, MNRAS, 441, 6
Seljak U., Makarov A., McDonald P., Trac H., 2006, Phys. Rev. Lett., 97, 191303
Shao S., Gao L., Theuns T., Frenk C. S., 2013, MNRAS, 430, 2346
Shaposhnikov M., Tkachev I., 2006, Phys. Lett. B, 639, 414
Shi X.-D., Fuller G. M., 1999, Phys. Rev. Lett., 82, 2832
Sigad Y., Kolatt T. S., Bullock J. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2000, preprint (arXiv:astro-ph/0005323)
Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760
Tikhonov A. V., Gottloeber S., Yepes G., Hoffman Y., 2009, MNRAS, 399, 1611
Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
Trujillo-Gomez S., Schneider A., Papastergis E., Reed D. S., Lake G., 2016, preprint (arXiv:1610.09335) (TG16)
Venumadhav T., Cyr-Racine F.-Y., Abazajian K. N., Hirata C. M., 2016, Phys. Rev. D, 94, 043515
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502
Villaescusa-Navarro F., Dalal N., 2011, J. Cosmol. Astropart. Phys., 1103, 024
Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740
Vogelsberger M., Zavala J., Cyr-Racine F.-Y., Pfrommer C., Bringmann T., Sigurdson K., 2016, MNRAS, 460, 1399
Wang J., White S. D. M., 2007, MNRAS, 380, 93
Williams L. L. R., Saha P., 2011, MNRAS, 415, 448
Yaryura C. Y., Helmi A., Abadi M. G., Starkenburg E., 2016, MNRAS, 457, 2415
Yoshida N., Springel V., White S. D. M., Tormen G., 2000, ApJ, 544, L87
Zavala J., Jing Y. P., Faltenbacher A., Yepes G., Hoffman Y., Gottlober S., Catinella B., 2009, ApJ, 700, 1779
Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, L20
Zwaan M. A., Meyer M. J., Staveley-Smith L., 2010, MNRAS, 403, 1969

APPENDIX A: COMPARISON WITH OTHER STUDIES

Several recent papers have reported solutions to the overabundance problem of the galactic VF within the standard model of ΛCDM. In this section, we discuss how these papers differ from our work and why we think they do not remove the problem.

Brook & Di Cintio (2015b) performed a detailed investigation of the galactic VF for three cases, CDM with baryonic cores, as well as one model of WDM and SIDM both without baryonic cores. Starting with the halo mass function, they use AM to obtain stellar masses, from which they estimate the extent of stellar and H I discs as well as baryon-induced DM cores. This allows them to obtain mock velocity profiles for each halo mass. Based on these velocity profiles, Brook & Di Cintio (2015b) determine the radius where the circular velocity equals the observed $v_{\rm rot}$ from small galaxies (which leads to an empirical relation between $M_{\rm star}$ and $r_{\rm rot}$). They find this radius to be much smaller than the radius of maximum circular velocity, concluding that $v_{\rm rot}$ is considerably smaller than $v_{\rm max}$. This bias between $v_{\rm rot}$ and $v_{\rm max}$ strongly reduces the initial discrepancy between theory and observation in the VF without, however, completely solving it. Brook & Di Cintio (2015b) show that the remaining tension can be solved by either assuming baryon-induced cores within ΛCDM or ADM (i.e. WDM or SIDM without baryonic cores).

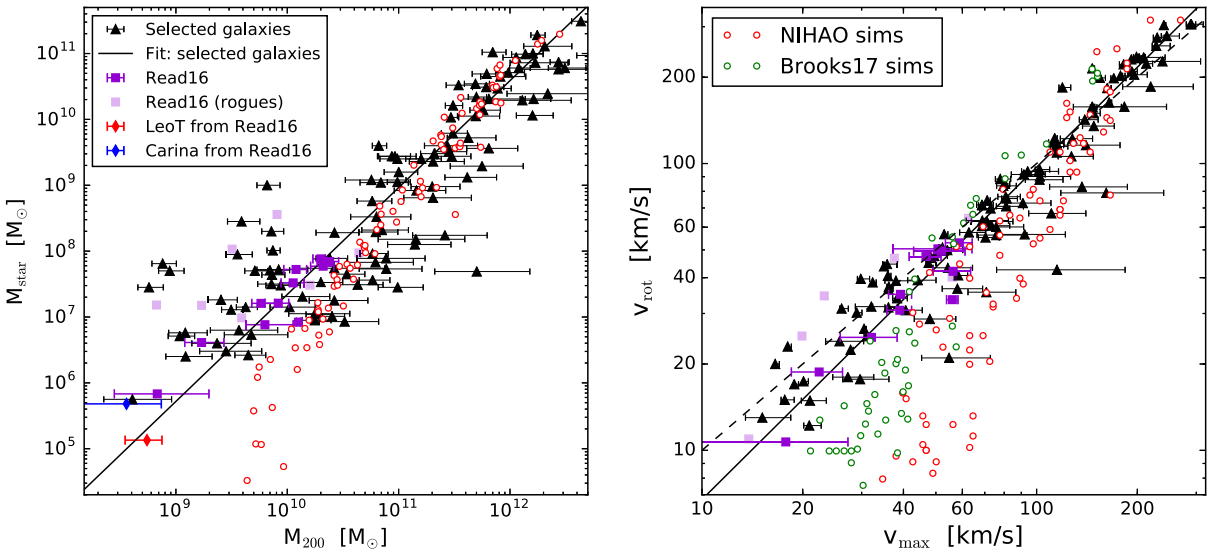

Figure A1. Relation between $M_{\rm star}$ and M_{200} (left-hand panel) and between $v_{\rm rot}$ and $v_{\rm max}$ (right-hand panel) for the selected galaxy sample (black triangles) compared to individual field dwarfs from Read16 (purple symbols). The bold purple symbols correspond to reliable rotation curves while the shaded purple symbols (without error bars) have potential systematics due to low inclinations or signs of disequilibrium (dubbed 'rogues' in Read16). The red and blue data points represent mass estimates for the LeoT dwarf and the Carina satellite (Read16). Next to the direct estimates from observations, we show results from the NIHAO and the Brooks17 simulations (red and green circles) which agree at large scales but deviate towards smaller scales most relevant for the VF.

The main weakness of the method applied by Brook & Di Cintio (2015b) is that it relies on abundance matching. The AM technique matches observations to the ΛCDM model by *assigning* small galaxies to very large haloes, without *testing* this assignment against observations. As a consequence, Brook & Di Cintio (2015b) obtain small H I radii (via their empirical $M_{\rm star}$–$r_{\rm rot}$ scaling relation) and therefore small values for $v_{\rm rot}$, which largely alleviates the initial discrepancy between the predicted and observed VF.

Very recently, Brook & Shankar (2016) published a paper, where abundance matching combined with the observed BTF relation is used to estimate the bias between $v_{\rm rot}$ and $v_{\rm max}$. They find that the discrepancy between observed and predicted VF disappears entirely, provided the adequate Tully–Fisher relation is used. While this paper reveals systematical differences between different observations and highlights the importance of an adequate velocity definition, it does not provide a test for ΛCDM on its own.[12] A further necessary requirement for testing the underlying cosmological model is to verify the applied AM relation.

Attempts to verify abundance matching of ΛCDM by estimating the halo mass of observed galaxies have been performed by several papers in the past. Papastergis et al. (2015) used kinematical information from the extended H I content of dwarf galaxies to show that there is a discrepancy between the data and AM relations from the local volume, irrespectively of whether baryon-induced cores are assumed or not. More recently, Pace (2016) estimated halo masses based on careful analysis of full H I rotation curves for a few field galaxies, obtaining similar results. Karukes & Salucci (2017) also find a discrepancy with AM expectations based on ΛCDM, by analysing the rotation curves of a sample of 36 late-type dwarfs. Brook & Di Cintio (2015a), on the other hand, used stellar kinematics to estimate halo masses of dwarf galaxies. They find better agreement with AM relations from the local volume, mainly due to their assumption of cored profiles from baryon processes. However, stellar kinematics only probe the very inner region of haloes, which are subject to large uncertainties and potential systematics in the halo mass estimates. Indeed, small differences in the model are amplified leading to large differences in halo mass.

Recently, Macciò et al. (2016) published a study on the VF based on the NIHAO (Numerical Investigation of a Hundred Astrophysical Objects) suite of hydrodynamical simulations with full metal cooling and standard recipes for subgrid effects such as star formation and supernova feedback. They report a very large bias between w_{50} and $v_{\rm max}$ fully solving the discrepancy between the observed and predicted VF. There are two main reasons why the results from Macciò et al. (2016) differ from ours. First, the H I content of the NIHAO galaxies is less extended than the one from the selected sample of observed galaxies leading to smaller values of w_{50} (or $v_{\rm rot}$) compared to $v_{\rm max}$ (see TG16 and Papastergis & Ponomareva (2016) for a detailed comparison). Second, the strong feedback effects present in the NIHAO simulations make the NIHAO galaxies reside in very massive haloes compared to the mass estimates from the selected sample.

Finally, during the review process of this work, Brooks et al. (2017, hereafter Brooks17) published a study based on a suite of hydrodynamical simulations that claims to fully solve the apparent discrepancy of the VF. Similarly to Macciò et al. (2016), they obtain a more significant bias between $v_{\rm rot}$ and $v_{\rm max}$ compared to what we find in our analysis. A closer look at their results reveals that they are able to completely close the gap between the observed and predicted VF between $v_{\rm rot} = 20$–50. However, there is some remaining discrepancy at both larger and smaller velocities, which they correct in the latter case by assuming a cut-off due to reionization.

As a result of very efficient feedback recipes, both the NIHAO and the Brooks17 simulations obtain larger halo masses and larger maximum circular velocities than what we find by analysing H I kinematics of dwarf galaxies. This is illustrated in Fig. A1, where

[12] Indeed, for the ideal case where the BTF, the VF, as well as the AM relation is based on one single set of observations, the discrepancy of the VF has to disappear by construction for a large number of different cosmological models.

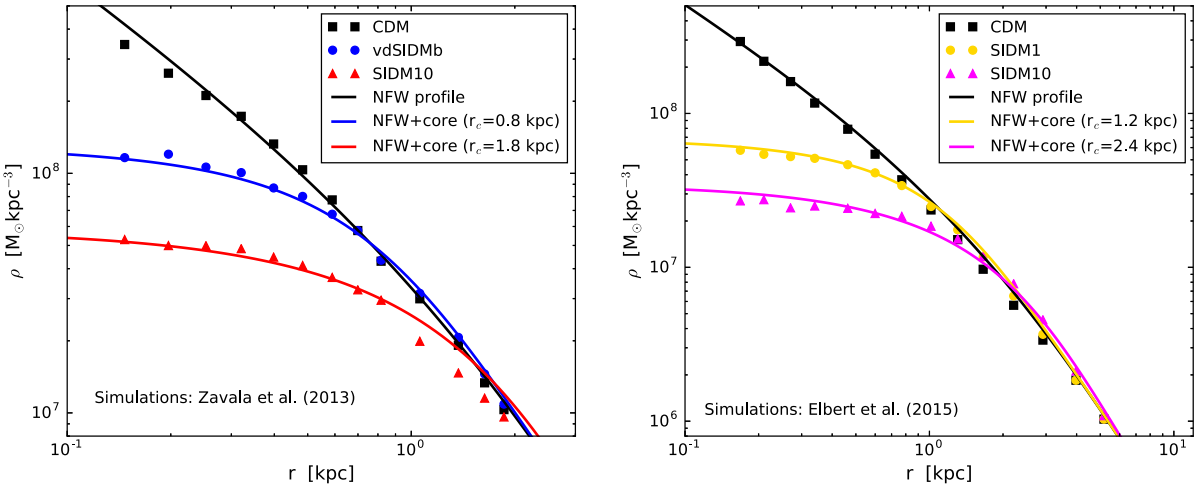

Figure B1. Simulated profiles of dwarf galaxy haloes from the literature (Zavala et al. 2013; Elbert et al. 2015) fitted with the R16 profile (coloured dots and lines, respectively). *Left:* the model vdSIDMb (blue) corresponds to velocity-dependent SIDM (with $\sigma_{\rm m}/m = 35$ cm^2 g^{-1} and $v_{\rm m} = 10$ km s^{-1}), while SIDM10 (red) refers to a velocity independent model with $\sigma/m = 10$ cm^2 g^{-1}. *Right:* two models with velocity independent cross-section of $\sigma/m = 1$ cm^2 g^{-1} (yellow) and $\sigma/m = 10$ cm^2 g^{-1} (magenta).

the left and right-hand panels show the $M_{\rm star}$–M_{200} and the $v_{\rm rot}$–$v_{\rm max}$ relations, respectively. For both relations the selected galaxy sample (black symbols with error bars representing the dependence on the concentration parameter) is well described by a power law (black solid lines), while the NIHAO and Brooks17 simulations (red and green circles) exhibit a downturn towards small mass and velocity scales. This is a direct consequence of their strong feedback recipes which reduce the amount of stars and gas in a halo of a given size.

In Fig. A1, we also compare our results to independent estimates of halo mass and maximum circular velocity from Read et al. (2016a, hereafter Read16). They used fully resolved rotation curves of individual field dwarfs accounting for stellar and gaseous components as well as baryon-induced cores. The results of Read16 are shown as purple squares, where bright symbols represent dwarfs with reliable rotation curves, while the shaded symbols without error bars denote data with potential systematics from inclination or signs of disequilibrium (dubbed 'rogues' in Read16). It is very encouraging that the results from Read16 agree well with our own mass estimates.

In summary, there is a systematic difference between the halo mass of simulated galaxies (from the NIHAO or the Brooks2017 simulations) and direct mass estimates from local field dwarfs which could point towards a genuine problem of hydrodynamical simulations at dwarf galaxy scales. However, we want to point out that the current observational data are too sparse to support any strong conclusions. Upcoming large area H I surveys with interferometric data will highly improve the observational situation in the next few years.

APPENDIX B: A NEW HALO DENSITY PROFILE FOR SIDM

Here, we demonstrate that the R16 profile (Read et al. 2016b) is not only suitable to describe cores from baryonic feedback, but also provides very accurate fits to profiles of SIDM haloes. In equation (1), we introduced the R16 mass profile which consists of the NFW mass profile multiplied with a simple two-parameter function. The density profile can be obtained from equation (1) by a simple derivative, i.e.

$$\rho_{\rm R16}(r) = \rho_{\rm nfw}(r) f^n + \frac{n f^{(n-1)}(1-f^2)}{4\pi r^2 r_{\rm c}} M_{\rm nfw}(r), \qquad \text{(B1)}$$

where $n = 1$ and $f(r)$ is given by equation (1). The R16 profile has the advantage of converging to the NFW profile for $r \gg r_{\rm c}$, where DM self-interactions are negligible.

In Fig. B1, we show fits using the R16 profile to a few simulated SIDM profiles found in the literature. The left-hand panel shows data points for two haloes from Zavala et al. (2013) corresponding to velocity-dependent model with $\sigma_{\rm m}/m = 35$ cm^2 g^{-1} and $v_{\rm m} = 10$ km s^{-1} (blue, dSIDMb) and a velocity-independent model with $\sigma/m = 10$ cm^2 g^{-1} (red, SIDM10). The haloes are well fitted with R16 profiles with $r_{\rm c} = 0.8$ and 1.8 kpc. The right-hand panel shows two haloes from Elbert et al. (2015) out of velocity-independent SIDM simulations with $\sigma/m = 1$ cm^2 g^{-1} (yellow, SIDM1) and $\sigma/m = 10$ cm^2 g^{-1} (magenta, SIDM10), respectively. Again the R16 profile provides an accurate fit to the simulated halo profiles. The core sizes for these haloes are $r_{\rm c} = 1.2$ and 2.4 kpc.

This paper has been typeset from a TEX/LATEX file prepared by the author.